Speaking to Reading in English

Investigating the Relationship between Prosody and Reading Comprehension with Year 3 Children

Sheradan Christopher James Miller

Supervisors: Dr. C. Dye
Dr. M. Jin
Dr. F. Smith
Dr. H. Pichler

School of English Language, Literature & Linguistics
Newcastle University
Doctor of Philosophy

April 2023

Abstract

Prosody is often described as the melody of language, and an important linguistic feature to comprehend speech. However, there is a lack of research investigating the relationship between prosody and reading comprehension with children learning to read. Prosody was measured with the PEPS-C 2015 (Peppé, 2015), and investigated in relationship with measures of reading comprehension. This relationship was also analysed with measures representing the aspects of The Simple View of Reading as control variables, which have been well-evidenced as predictors of reading comprehension in prior research. The project worked with 51 children aged between seven and eight in a primary school in the North-East of England. The results suggested that in agreement with previous research prosody does contribute to reading comprehension ability, though this relationship is altered by the presence of the components of the SVR. The exploratory simple mediation analyses further reinforced this. We concluded that children's understanding and use of prosody in our study was predictive of reading comprehension ability on the tasks. The current project is limited by the cross-sectional design, making conclusions about the causal impact of prosody on reading comprehension not feasible at this time. Further research should follow the relationship between prosody and reading comprehension through a longitudinal study, and could utilise alternative methods (e.g. eye-tracking) to assess more fine-grained aspects of this relationship.

Dedicated to Arthur Robert Noel Harker and Dr Cristina Dye . . .

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 100,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

Sheradan Christopher James Miller
April 2023

Acknowledgements

This is the self-indulgent section, where I get to shed the academic front and simply get to be appreciative of the people in my life who helped me through this project. There are so many people that have contributed in so many different ways to complete this project. Whether through emotional support, supervisory guidance, anonymous guidance with LaTeX templates or R code, those who listened to me at conferences, and so on.

Firstly, I would to acknowledge my supervisors role in the completion of the thesis; Dr Cristina Dye, Dr Michael Jin and Dr Faye Smith. They guided me through the hoops of being a PhD student/researcher, listened to me at my highest and lowest in talking about the project, and helped me towards completion. Furthermore, I would like to thank Dr Simon Gibbs, who helped me at the start of the project with good probing questions and nice green tea in his office. I'd also like to thank Dr Kim Pearce, who went out of her way to help me with my questions regarding statistics for the project when she didn't have to, and taught me so much in a short time that helped a lot. There are so many other staff members whom I met with only once or twice, but through discussion made me think about my project critically and changed my mind, they all helped me more than they will ever know.

I would also like to acknowledge my writing group buddies - Caitlin, Andreas, Joiquin, Jasmine, Bella, Olivia, Damar, and Fengting who helped me to write throughout the final few years of the project. Having a group of intelligent and engaged PhD researchers to ask questions and listen to was a great help, and despite being a socially awkward person I felt like I was part of our own little academic club that got me through the hardest writing blocks. I would also like to thank an online supporter, Tuna, on a forum I follow. It's nice that even through pretty much anonymous forum interactions I was able to receive support from someone who was also going through the same experience, thanks a lot! Furthermore, I would like to thank my friend Olivia for being a good listener to my research quibbles and doubts, and for lifting me up when I was down. I hope your own PhD is successful! As well, my friend Jay has been a continuous support throughout these four or five years, with extroverted passion to help this introverted guy get through, thanks so much!

Related to my thesis, it would not look as neat and tidy as it does without some amazing previous PhD students and their LaTeX expertise. Krishna Kumar created the original PhD template and posted it via Github, it is such a great piece of coding that took away so much of a headache so thank you. Furthermore, Newcastle computing science student Matt Forshaw designed the initial template for a Newcastle University thesis, which was expanded on by

Table of Contents

List of Figures

List of Tables

Glossary

alphabetic principle

The understanding that the written word represents the spoken sounds of the language. That is, the graphemes of the writing system are representative of the phonemes of the language.

expressive prosody

The production of prosody, how prosody is used in speech by the speaker. In the current project this related to the PEPS-C 2015 tasks which required the participant to say text aloud using the appropriate prosody.

implicit prosody

The idea that readers impart prosody onto what they are reading to understand the message intended by the writer. First well known discussion of this phenomenon was by Janet Dean Fodor (1998, 2002), who describes that to garner understanding from a text readers must provide the prosody intended by the writer to gain the intended message for themselves.

linguistic comprehension

The ability to understand and derive meaning from spoken language. The ability to extract and construct meaning from the speech stream. Often described in parallel to reading comprehension, with the only difference being the representation of the information. It is a complex construct, involving skills such as working memory, vocabulary knowledge, grammatical knowledge, attentional control, and more. It is one of the two predictors of reading comprehension in the Simple View of Reading Framework.

phonics

A method of reading instruction used in UK schools. The intention is that children are taught to understand the relationship between the graphemes of written language and the phonemes of spoken language. The hope is that children will learn the correspondences between these aspects of language, and that it will help to independently read and write their native language.

phonological awareness

The awareness and understanding of the phonological structure of words.

prosody

Also known as suprasegmental phonology, prosody is an aspect of speech that adds meaning to the words being spoken. More specifically, prosody is can be defined by speech features such as rhythm, correct stressing of works, the intonation used to present speech. In relation to reading, prosody aids reading fluency, allowing children to read with accuracy and speed, and conveying the message of the written text to a listener. “How something is said rather than what it is" (Veenendaal et al., 2014).

reading comprehension

The ability to read text and interpret and construct meaning from what has been read. This can span from a literal interpretation of events which are occurring in the text, to inferences based upon the context and language used. Reading comprehension requires many skills such as strong decoding skills, the ability to infer hidden meaning, and others working in tandem to work effectively.

reading fluency

The ability to read text with speed, accuracy and accurate expression for what is being read. Reading fluency is a developmental skill, where children go from stop-start reading, to reading text with a voice like natural speech.

receptive prosody

The understanding or awareness of prosody use in speech. In the current study it refers to the tasks in the PEPS-C 2015 which required listening to speech, and being able to comprehend prosodic use.

Acronyms

CFF

Cognitive Foundations Framework.

PEPS-C

Profiling Elements of Prosody in Speech-Communication.

SVR

The Simple View of Reading.

TOWRE-2

Test of Word Reading Efficiency - Second Edition.

TOWRE-2 PDE

TOWRE-2: Phonemic Decoding Efficiency.

TOWRE-2 SWE

TOWRE-2: Sight Word Efficiency.

TROG-2

Test for Reception of Grammar Version 2.

YARC

York Assessment of Reading for Comprehension.

YARC SWR

YARC Short Word Reading Test.

Introduction

Prosody is sometimes referred to as the "music of language" (Wennerstrom, 2001, p. 14) or the "tune and rhythm of speech" (Carlson, 2009). These high level descriptions prescribe an importance to prosody's use in speech, and the current project investigated the relationship between prosody and reading comprehension. A key framework in the teaching of reading in the UK is The Simple View of Reading (Gough and Tunmer, 1986; Hoover and Gough, 1990), which posits that the two key components to becoming successful at reading comprehension are decoding (understanding the relationship between speech and the written word) and linguistic comprehension (understanding meaning in speech). The SVR is an influential framework on UK Educational Policy regarding the teaching of literacy, beginning first with its recommendation in the Rose Report (Rose, 2006). However, the SVR does not consider the explicit influence of prosody on reading ability, a component of speech that gives extra meaning and clarity to speech above and beyond the words spoken themselves. This thesis was an exploration of the relationship between children's understanding and use of prosody and their reading comprehension ability, while controlling for the components of the SVR framework.

Chapter 1. Learning to Read: From Speech to Text

The importance of learning to read cannot be overstated in the modern world. People can live a fulfilled life without literacy, but it makes one's life much more difficult to navigate. There is an assumption that people are literate in society, from signs in the street, to emails, to reading a recipe to eat, all this requires a basic level of literacy. Beyond this, literacy is important for social and economic opportunities (Thomson and Jarmulowicz, 2016, p. vii), the majority of work requires an assumed level of reading ability, and if lacking, may affect a person's future living conditions. Furthermore, low literacy skills can be a major determinant of inequality in society (Castles et al., 2018), so it is important from an early age that children are given the best chance to access and develop their literacy skills, to best provide them with an advantage in the world.

However, literacy should not only be thought of as a method by which to prepare the young for their working lives. Reading ability also leads to self-teaching and creative endeavours, that enrich our lives and ability to engage with the world. Imbuing a love of reading with children from an early age can help them to explore their interests, whether through fantastical fiction, or details about the animal kingdom from a non-fiction book. One of the main ideas I express throughout this thesis is the endpoint of reading being the ability to read independently for meaning, to take the written word and interpret meaning for personal knowledge and satisfaction.

Prior to further discussion of reading development and the linguistic components of the current thesis, it is useful to define what the term reading means in the current thesis. At its most basic, reading begins as the ability to look at a piece of written text and understand that the written words represent the sound structure of a spoken language. Taking the English language as an example, we come to recognise that certain combinations of letters make certain sounds, that is linking the visual symbols to their corresponding phonemes (Ziegler and Goswami, 2005). As the act of reading beyond word-level becomes automatised, what one gains from reading and what they are able to read becomes more complex. Meaning needs to be attained from the sentences and paragraphs being read, which involves a variety of linguistic and cognitive skills to be activated and working effectively in tandem (Kim et al., 2018). The goal of reading is to be able to decode the written language and construct meaning from this text (Castles et al., 2018; García and Cain, 2014).

Reading is not an innate skill, and requires time and effort for the learner to acquire. Furthermore, the time and effort provided by those who teach reading (school teachers, teaching assistants, parents) is also of major importance, for it is them who will instil a love of and the ability to read. It is thus of two-fold importance that the scientific community continues to investigate and understand the science of reading, to aid the educational system in the ongoing understanding of what helps humans to read. There is a discrepancy between the understanding of the academic world and the classroom about the science of reading (Castles et al., 2018), with educational policy behind where the reading science field's knowledge currently is. However, this may be due to a lack of communication between the academics and the field of teaching, a failure on the academic's part to communicate such findings in a suitable way to aid teachers work.

There is also a relationship between spoken language and written language. But, reading is not the natural next step from speech in the development of language. Children have to go through the drawn-out process of coming to understand their own language in written form. The same linguistic features are present in text as they are in speech; the sounds represented by the words in text (phonology), the differing spellings of similar sounding words (morphology), the placement of stress and pauses in sentences affecting meaning (prosody), the meaning of the words in context (semantics), and how the words and phrases are arranged (syntax). However, there are differences between speech and writing that go beyond linguistic factors. For example, speech is both spoken and heard which requires speakers to be clear with the language spoken. Furthermore speech is influenced by the speakers use of emotion to share their message, which shapes the meaning being expressed.

What is important to remember is that most children do become skilled readers. Though some learners may leave school and find they lack the want to read a book for personal pleasure, they will still be literate and have a skill that is vital in the modern world.

1.1. The Framework of a Successful Reader

Reflecting on the importance of reading to children and their future life success, the next step is to discuss how one comes to read successfully. Reading is a skill made up of many component skills, that attempting to cover all this in a single model or framework is hard to comprehend. As such, focus is put upon certain aspects of reading development, and to build up understanding of how the particular skills make a reader. Particularly in understanding how children proceed to successful reading comprehension is a challenge, as it involves many components, such as successful word reading, the ability to monitor and retain meaning from text, differentiate genres of text and appreciate context, competent eye movement ability, fluency of reading etc. Understanding and determining what makes great reading comprehension skill is a field in and of itself, and is not a major focus of the current thesis, but suffice to say there is a reason the field of reading science continues to be fruitful for exploration and understanding.

Theories of Word Reading

Theories and frameworks have attempted to provide comprehensible ways of understanding how readers develop to be successful. Many of the examples for early readers discuss the initial development of word reading, a key starting point for starting to read independently. A popular example is by Ehri, who popularised the theory of how initial word reading develops (Ehri, 1995, 1999, 2014). In short, children develop word reading skills from an initial understanding that the printed word represents elements of their spoken language, but do not yet have the context to read and rely on their visual memory to do this (pre-alphabetic phase). This progresses through a stage where children begin to show knowledge of letter-sound correspondences (partial alphabetic phase), to the ability to decode new words with some effort but rules and regularities of written language have started to internalise (full alphabetic phase), ending at a point of ever-maturing ability wherein continued exposure to writing provides automatised and effortless reading ability to occur (consolidated alphabetic phase). This represents the early development of reading ability, and is an attempt to reflect the complexities children are going through to even just come to read words, never mind comprehending their meaning in context.

Another important theory of reading is by Share, The Self-Teaching Hypothesis (Share, 1995, 1999). This model of reading acquisition focuses upon the importance of decoding in early reading development. The idea is that with each successful decoding of a new word during the independent reading of a text, this presents the child an opportunity to create a specific print-to-meaning connection. This is phonological recoding, and this self-teaching function provides the reader with the ability to progress in reading by reading more and more, and continually developing these print-meaning relationships. The idea of self-learner is a pertinent one, as it provides an explanation that early-stage readers can develop the knowledge to teach themselves reading early on, and explains why Share's model has maintained discussion in the science of reading.

Each of the above theories are focused on word reading specifically, which is an incredibly important part of early reading development. Both are similar, in that they focus on a developmental view of word reading, that success develops over a space of time, and becomes internalised. They both provide an account of what successful orthographic learning is, that is to read words with an automaticity that frees up cognition for other tasks. The weakness of both of the approaches however, is an explicit explanation of how children to this outcome.

Theories of Reading Comprehension

When you move further from word reading to then reduce reading comprehension to a single framework or theory, it becomes more complicated. In the introduction to their reflection on reading comprehension theories, Perfetti and Stafura (2014) open with the bold statement that "there is no theory of reading", the idea being that reading is composed of too many individual components to reasonably encompass all of its many quirks and complexities. Indeed, Perfetti themselves designed a model of reading ability named the Lexical Quality Hypothesis (Perfetti and Hart, 2002) (henceforth referred to as the LQH). In short, this theory presents the idea that successful comprehension is dependent upon "accessible, well-specified and flexible knowledge of word forms and meanings" (Perfetti and Adlof, 2012). The model is made of two components, the orthographic system (orthography is the conventions for representing a language in writing, with children developing their orthographic reading skills to develop word recognition) and the linguistic system, with both levels interacting with each other to develop successful reading ability. In essence, the better the quality of lexical representations, the more successful children will be with spoken and written language. This is another theory of reading that could have an entire thesis written to explore, but I present it now as a contrast to the framework used for the current thesis. Perfetti and Hart focus on the interaction between two components of reading, and how through cognitive resources working in tandem can lead to successful text comprehension.

The construction-integration model (Kintsch, 1988) focuses on a more multilevel approach to reading comprehension. The model describes comprehension as having two phases that help the reader to obtain new levels of text understanding. The first is the construction phase, a bottom-up process that allows the read to build an initial representation of the text from their initial reading. This *surface code* code level, as described by Kintsch, captures the exact wording and syntax of the text, and is then replaced by the second phase, the *textbase* which captures the basic ideas in the text. This is developed through a set of *propositions* that can cause the reader to retrieve information from their long-term memory that may be associated with the text (e.g. Ann is a name, therefore she is the subject of the text). The model captures an aspect of comprehension that it is not just the lexical aspect of reading that is important, but also external factors such as prior knowledge that can help an individual build meaning. This helps predict why readers should be better with comprehending a text they have prior knowledge about (Schwanenflugel and Klapp, 2015, p. 183). Furthermore, the model can explain why some texts take longer to read than others, as they may be denser in the aforementioned *propositions*. For the current project this model was not used, similarly to Perfetti and Hart above, as this focuses more on the integration of knowledge into comprehension, rather than being a breakdown of the specific components of what makes successful reading comprehension.

Another example is the schema-theoretic view of reading. Described by Schwanenflugel and Klapp (2015) in their chapter on reading comprehension theories, they define schemes as "a knowledge structure that organises and summarises what a person knows about a topic" (pg.172). Developed by Anderson (1994) as way to explore reading comprehension, this model suggests that the schemas affect what readers learn and remember from a text. As such, when reading a text readers begin searching for the information they know which can help them to make sense of a text. To be able to locate an appropriate schema has a direct correlation to the reader's ability to comprehend text (Schwanenflugel and Klapp, 2015, p. 175). This is a neat simple model, and has history in cognitive psychology's exploration of schemes. However, for the current study it was not suitable for adapting to include the aspect of language to be explored in this study, prosody. This was the common goal when deciding on a model to design the methodology around, and will be discussed in more detail in Chapter 2.

Finally, the direct and inferential mediation model of reading comprehension (DIME) by Cromley and Azevedo (2007). This model is closer to what this project was searching for as a model, as it details the components of language which contribute to successful comprehension. This is presented in a model, as seen in Figure 1.1, with several variables listed:

- Background - background knowledge on a the text being read
- Strategies - awareness and use of strategies readers use to comprehend text
- Vocabulary - the breadth of vocabulary a reader has
- Word - word reading ability
- Inference - the ability to infer meaning from text
- Comprehension - the outcome measure for the model, the ability to comprehend meaning from the text

The interesting aspect of the DIME model is that it attempts to integrate components from text processing theories of comprehension (Ahmed et al., 2016), such as background knowledge, into the more common cognitive measures in reading models, like word reading and vocabulary. The model is also a useful design for creating a structural equation model, a complex but detailed modelling measure that allows variables to be entered into the model and direct and indirect effects measured. An example of this analysis was conducted by Cromley et al. (2010), who wanted to test the validity of their model by applying to a domain specific subject, that being biology students at university. They found that their domain-general model did show an excellent fit with the domain-specific biology text they issued with the students, suggesting suitability beyond generalised texts. Similarly, Ahmed et al. (2016) wanted to test the validity of the DIME model with students in grades 7-12, to see if the model was a good fit with this age group. In summary, the researchers found that the model accounted for the majority of systematic variance in reading comprehension, that is the variables in the model accounted for most of the reading comprehension ability of students. They found that the importance of knowledge and vocabulary was diminished over time, and that inferencing became more important in the later grades. Again, the authors suggested the DIME was a suitable model for generalising to measure reading comprehension.

This overview of various theories and frameworks is to bring attention to the important role of spoken language in reading development, and to provide examples of different approaches to modelling this relationship. The commonality is in the importance of children internalising the rules of their language, and becoming more automatised in their ability to be more efficient decoders and comprehenders of language. However, the current project did not use any of the theories and models discussed in this section. This was due to many issues, for example the DIME model is well evidenced, but for the scale of the project to be undertaken it was thought not enough data could be collected for each variable to have the statistical strength for the model

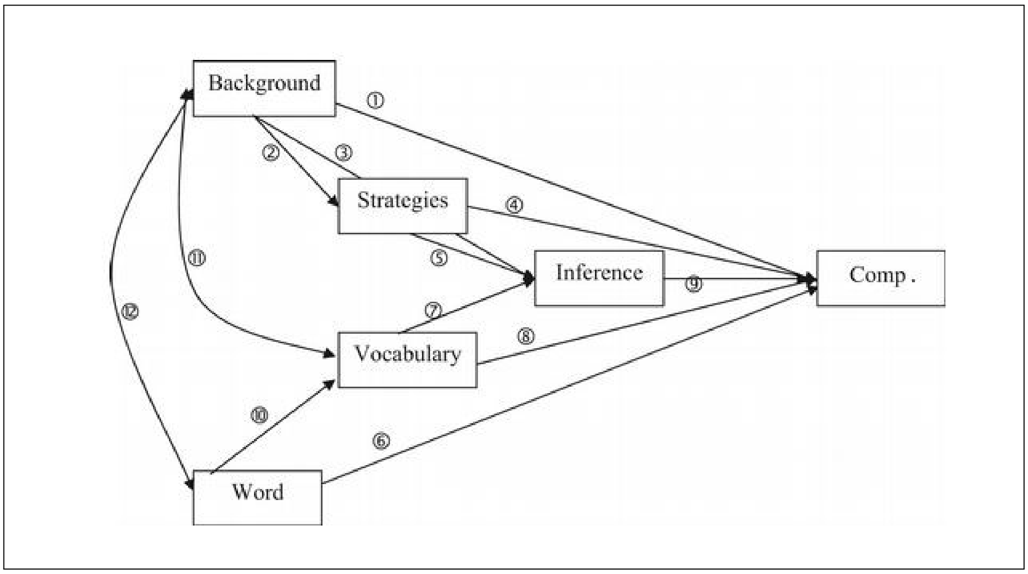

Figure 1.1 DIME model of reading comprehension

to be a good reflection of reading comprehension ability. In the next section we will explore the framework chosen for the project, the Simple View of Reading.

The Current Study and Choice of Framework

It should be clear from these initial descriptions that there are many ideas about the development of reading ability in children. Each, though different in the process children go through to read, all rely on the fundamental idea that children's familiarity and experience with word recognition and decoding is both a prerequisite and an aid to later reading ability. Furthermore, the letter-sound relationship is reflected in these descriptions, that knowledge of the spoken word is an influence upon literacy. There is a framework of reading development that has not been mentioned yet which has had a major influence upon reading development thinking, such that it affected the UK Educational Curriculum for how reading is taught (Rose, 2006).

1.1.1. The Simple View of Reading

One of the many ideas of learning to read is The Simple View of Reading (SVR). Developed initially by Gough and Tunmer (1986), the SVR attempts to break down what leads to successful reading comprehension into two components. Reading comprehension can be seen as the endpoint of reading development, as with its acquisition reading becomes self-directed and individualised to allow those with the skill to understand and be critical of the texts they engage with. This differs from just word reading, which is the ability to sound-out and produce the sound of written language, it involves understanding what is being read in the moment. Gough and Tunmer's suggestion was a framework which was made up of just two components that they believed would lead to success in reading comprehension. These two components are *decoding* and *linguistic comprehension*. The two components are displayed as part of an equation, a simple way to represent the development of reading, shown in Figure 1.2.

$$\boxed{\text{D x LC} = \text{RC}}$$

Figure 1.2 The Simple View of Reading equation. D = Decoding, LC = Linguistic Comprehension, and RC = Reading Comprehension.

Reading comprehension is presented as the product of decoding and linguistic comprehension. It is important that it is presented as a multiplication, indicating that an improvement of either aspect of these components is beneficial to the overall development of reading comprehension.

The intention of the SVR in the original paper by Gough and Tunmer, was an attempt to explain why different reading disorders occur, and how they could be represented by this framework. One example is dyslexia, which is marked by difficulty with decoding words, being able to translate the written print to its spoken form i.e. correctly pronouncing the written words. To note, this is a major simplification of dyslexia and its complexity, but for the purposes of the SVR model dyslexia is simplified into this definition. Difficulty with decoding is a component of dyslexia and may affect the ability to reading, and in terms of the SVR model it is important that children's language comprehension is not affected. The authors highlight that children may be perfectly fine with understanding and using their spoken language, but when asked to read the words on the page the lack of decoding ability hampers them garnering meaning from text. This is the purpose of the multiplication sign in the above equation, each value of D and LC ranges from a theoretical 0 (no knowledge) and 1 (perfection) in Gough and Tunmer's design. Decoding is very rarely going to be 0, even in dyslexia having no ability to decode is incredibly rare, but decoding ability will be substantially lower for children with dyslexia. This idea was expanded upon in a follow-up paper by Hoover and Gough (1990), an assessment of the SVR in a research context with a group of English-Spanish bilingual children from first through to fourth grade. Their results provided further support to the framework, decoding and listening comprehension contributing to variation in reading comprehension when measured with linear regression.

The SVR definition was returned to by two of these authors (Hoover and Tunmer, 2018) in a special issue of the Remedial and Special Education journal. They reflected on the SVR and its utility after 30 years. Some important notes they add to the discourse about their framework and its definitions was to say that though the SVR separates the complexity of reading into two component parts, they do not claim that reading is simple. There is a clear rebuttal to be made, and one which I will show in Section 1.1.1 when discussing critique of the SVR, which is why is it named the Simple View of Reading? The argument is that the framework is simple, but the components representing reading are not. It is easy to remember and a powerful conceptualisation, as it relies solely on basic mathematic skills to comprehend (Uppstad and

Solheim, 2011). We will see below as I focus in on each of these components that there is more complexity within their definitions than it first appears.

Another point Hoover and Tunmer highlight is that the SVR is *not* a model of reading development, rather it is a framework to describe reading at a single point in time. The multiplicative combination of the two components is the important aspect of how reading comprehension occurs. Early on, decoding is the most important contributor to reading comprehension outcome as children develop their word recognition ability through systematic phonics, developing phonemic awareness so decoding becomes automatic (Foorman et al., 2015; Verhoeven and van Leeuwe, 2008). Once this fluency in decoding is reached, then more attention can be paid by the readers towards comprehension, which is where the child's ability to comprehend speech becomes more important, including aspects such as their vocabulary. The shifting relationship of the two components to reading comprehension is important to remember, and is not necessarily something clear from the framework itself without this extra information given by the authors. Rather, decoding and linguistic comprehension differ in importance for each individual child dependent on their current reading ability, age, and where they are in their education.

Next, I move focus to each of three aspects of the SVR framework in turn, discussing how the authors defined these components and what they can potentially tell us about what makes a successful reader.

Decoding

The first component of the SVR is decoding. Decoding is a foundational skill in early reading development, providing the basic skills for readers to become successful (Van Wingerden et al., 2018). At its most basic, decoding is the ability to recognise and utilise the knowledge that the written characters on the page represent the spoken sounds of language i.e. correctly recode the written words into a spoken form. Experienced readers take for granted that the word 'cat' is made up of three letters, three phonemes (/k/, /a/, and /t/), and that this particular combination of phonemes makes the word 'cat'. Becoming a proficient reader means that this aspect of reading is automatised, we do not notice that we are processing these words so fast. However, when a child begins developing reading ability learning to decode is quite difficult.

This is particularly important for the language being explored in this thesis, the English language. Despite being a language that is explored and researched to an incredible amount, the English language is not particularly easy to decode (Lyytinen et al., 2015; Seymour et al., 2003), both for second language learners or indeed younger first language speaking children. This is because the English language is quite an opaque language, the relationship between the letters and sounds is not one-to-one. This is more clearly highlighted when thinking about how there are 26 letters in the English alphabet, but spoken English has more than 40 phonemes in which to be represented by this alphabet. This current thesis does not have the space to provide a full historical context for why this is, what is important is that developing proficiency in English decoding skills is difficult and requires years of good teaching and practice.

Decoding is foundational, it grants readers the ability to decode words they do not already know how to say. A child encountering ‘bat’ for the first time will have an easier time decoding this word if they have previously encountered and understood ‘cat’, as they will simply have to swap out /k/ for /b/ and they can say this new word. It is important to recognise that decoding does not mean the child comprehends the meaning of the word simply by reading it correctly, rather it provides a means of cracking the language’s code to allow passage to comprehension.

The development of decoding ability is important as it constrains reading early in development (Castles et al., 2018). This constraint is the effort it takes to decode each phoneme in a word, and then to put these sounds together (known in phonics as blending) takes time and mental effort for young children. This becomes more automatic over time through repeated exposure to words and appropriate phonics teaching in schools (Castles and Nation, 2008; Kuhn et al., 2010; Nation, 2009). This is an early level of reading fluency, the ability to read fast and accurately (Schrauben, 2010). With this fluency comes the ability to focus more on reading for meaning, because the cognitive resources that were being used to concentrate on decoding can now be utilised for meaning-making (Paige et al., 2012; Veenendaal et al., 2015).

As children grow older it has been suggested that the predictive relationship between decoding and reading comprehension ability weakens (Bosse, 2015). As discussed, the SVR was designed as a static model to describe reading at a single point in time (Hoover and Tunmer, 2018), and as such the changes in contribution by the two key skills can change and develop over time. What becomes more important are the skills that contribute to understanding the meaning in the text, such as semantics, syntax and pragmatics; the skills which go beyond simply identifying what the words sound like on the page to their greater meaning and purpose in a text. Decoding is a foundational skill for later reading ability under the SVR framework, but it alone cannot make a good reader.

Linguistic comprehension

In discussing decoding, I focused upon the ability of children to recode written words into spoken form. The wide breadth of this skill has been studied a lot throughout the literature, but less well-studied is the other contributing component to reading comprehension, linguistic comprehension.

Linguistic comprehension, as its name suggests and critics of the framework point out, is a phrase that can have a very wide definition. Hoover and Gough (1990) describe linguistic comprehension as a broad construct themselves, one quote taken from this paper is that linguistic comprehension is "the ability to take lexical information...and derive sentence and discourse interpretations" (p.131). An updated definition from two of the original authors refers to defining linguistic comprehension in parallel with reading comprehension, the only difference being the mode of representation of the information (Hoover and Tunmer, 2018). A potentially controversial comparison to make, the authors compare the two by explaining linguistic comprehension is about the ability to extract and construct inferred meaning from speech, which is the goal of reading comprehension with written text. One could counter that linguistic comprehension differs from reading comprehension in more than just the representation of information. For example, spoken language can be affected by who is speaking, However, this is part of the framework's design, that this early experience and expertise will contribute to the development of reading comprehension ability, and hopefully these skills are transferable towards meaning-making.

I mentioned linguistic comprehension has received less attention in primary education compared to decoding. This may be due to the complexity of the component and of itself, even without considering its relationship to the other components of the SVR (Kim, 2017). Linguistic comprehension skills can involve working memory, inhibitory control, vocabulary knowledge, grammatical knowledge, inferential ability, perspective taking, and comprehension monitoring (Cain, 2009; Florit et al., 2019; Kim, 2015). This is likely not even an exhaustive list of skills probed by the linguistic comprehension tasks I will discuss soon. However, the definition of linguistic comprehension does not ignore this complexity, it may be that this wide definition is an attempt to account for all these varying skills which contribute to a successful comprehender of speech.

Furthermore, it is important to differentiate linguistic comprehension from decoding as the other aspect of this equation. It requires more than just knowing how the words sound and the combination of these sounds. Even more than knowing the meaning of individual words, it is the ability to put all of this knowledge together in context to understand speech. Listening to speech is an active activity, which requires the individual to practice conscious processing of the speech stream. Furthermore, other higher-order skills such as comprehension monitoring and inference-making could be important to understanding (Kim, 2015; Kim and Petscher, 2016).

As will be discussed below, critiques are made of the SVR because of the generalisation of concepts, and linguistic comprehension is a major aspect that can be interpreted in many ways. However, as with decoding, the interpretative nature of the definition allows for an exploration of various complex aspects of linguistic comprehension.

Linguistic comprehension is the name given to this component in the original equation, but it has also been known as language comprehension and listening comprehension in the literature. Of particular interest is the use of listening comprehension, which has become an interchangeable term because listening comprehension tasks are a widely used measure of linguistic comprehension in reading science literature (Wolf et al., 2019). Linguistic comprehension has been operationalised in a number of ways, whether this is vocabulary tasks, listening comprehension, or combinations and alternatives (Kim, 2017).

One such example of a listening comprehension task is the Test for Receptive of Grammar 2nd Edition (TROG-2) designed by Bishop (2003). This task involves children listening to short sentences spoken by the researcher, which have varying grammatical constructs. The participant listens to the researcher read a sentence, and then points to a picture from a choice of four they feel best represents the sentence they just heard. One example is *The sheep the girl looks at is running*, requiring the child to listen carefully to ensure they comprehend that it is the girl who is looking at the running sheep, and choose to appropriate picture. This taps into the skills mentioned above, fast processing of the speech stream, comprehension monitoring, understanding of the vocabulary used. This task provides us with an understanding of linguistic comprehension as a relationship between how a listener is able to both comprehend a sentence, but also use this information to complete a task. It reflects the need to listen to instructions in the real world, and to use this information practically. The TROG-2 is useful in representing linguistic comprehension as it requires children to use their ability to monitor a sentence, recognise the use of grammar, and then provide an active response to this with the pictures. Interestingly, Whalley (2017) used this measure in a unique manner. They created a reading comprehension variant which matched the sentences, but had them written on the page for oral reading and was counter-balanced by swapping the nouns in the sentences where this made sense. This in theory made the only difference between the two measurements the presentation of the information (heard speech vs. oral reading), and was hopefully a matching task. This is an interesting reflection of the SVR, that the goal of comprehension is the same for these skills, but that they exist in different modalities. This may be too simplistic a difference, as listening and reading are more than just two different presentations of information. But when conducting research within this framework it provides a way to differentiate between these skills.

An example of the vocabulary measure of linguistic comprehension is the Peabody Picture Vocabulary Test-3 (McKinlay, 2011), in which a speaker reads off a list of words and the participant selects from four pictures which corresponds to the target word. This is strikingly similar to the above task, and this may be because when working with younger children tasks are designed for easy understanding, so in this case a receptive design wherein children simply have to point at a picture they recognise. A British version of this test is seen with the British Picture Vocabulary Scale, which follows the same procedure as above. Again, this task is an attempt to capture linguistic comprehension with a simple method whilst allowing children to hopefully use their linguistic skills to correctly identify the right picture.

A final example of a linguistic comprehension measure is the Test of Narrative Language (TNL) developed by Gillam and Pearson (2004). This task measures various aspects of narrative understanding, such as listening to an individual read a narrative text and to then answer comprehension questions, or to be given a picture of a scene and to create a narrative verbally. The narrative listening task is of particular interest, as a comparative point to reading comprehension measures (discussed below), which can also be measured by answers to comprehension questions. Compared to the previous examples, the TNL requires children to maintain their attention for a longer period of time. This is an extra challenge for developing learners to consider, and helps us to understand that linguistic comprehension is not just the ability to comprehend meaning, but the ability to maintain focused attention for a purpose. This is natural in everyday conversation, but this task tries to test this skill in isolation, seeing how well the participant can internalise information whilst maintaining attention.

Such a list of tasks for linguistic comprehension could continue onwards, but highlighting the three types that have been used above there is a difference in text type/genre (single word, sentences, short-form narrative) and participant response (receptive selection of picture, answering questions about a text). The choice of task could change depending on your participants, the type of text you want to explore, it depends on how a researcher decides to define linguistic comprehension. It is important to think about the choice of task though, and not just choose randomly. As Wolf et al. (2019) discuss, each type of comprehension is reliant on some modality-specific skills to process information. For example, linguistic comprehension is dependent upon listeners paying attention and storing information in their memory, as they cannot re-check heard speech as it is gone. Compare this to reading comprehension, where the text is there to check if the reader feels they missed something. This is what Kim (2017) means when she says that linguistic comprehension is a complicated construct in and of itself, even without thinking about how it fits into the SVR.

In their paper critiquing the SVR, Uppstad and Solheim (2011) highlight a specific case of how linguistic comprehension has been measured. They believe that the majority of measures are not speech as it would be normally encountered, that is the methods have individuals reading text aloud to a person, or listening to recorded speech. Either way, this is not the same as a normal conversation, with the unique quirks of individual discourse and other aspects such as body language. This is a more general issue with creating behavioural and cognitive measures which do emphasise natural behaviour. It may be more pertinent to focus upon the variety of context in measures instead. For example, the differences between a long-form story versus listening to a sentence, between listening to non-fiction versus non-fiction, a natural conversation versus an interview.

Reading comprehension

The final aspect of the SVR to discuss is what may be regarded as the aim of learning to read, the ability to read text and comprehend meaning. Reading Comprehension is one of the most important transferable skills for achievement in both education and society at large (Calet et al., 2019). Reading comprehension requires children to quickly and accurately recognise the words in text (decoding), while also constructing meaning from these words in the context of the text (Groen et al., 2019). To be able to decode the written word is the beginning of reading, but simply being able to understand the combination of phonemes and what this says is not enough. It is then that linguistic comprehension can aid us, knowing the meaning of the decoded word puts into context. Decoding and comprehending each word automatically, reading with fluency and understanding, this then materialises in reading comprehension. This is where the simplicity of the SVR framework materialises, the interface between decoding and linguistic comprehension brings a reader towards successful reading comprehension.

The foundation of reading comprehension comes from spoken language, but to become a successful reader of written language requires more than this. Beyond being the extended use of the speech-language system (Ashby, 2016), reading comprehension requires readers to bring their knowledge of other aspects of language into use. This is where the decoding aspect of the SVR can be highlighted, as children bring their knowledge of phonological information to the text, that is their knowledge of the speech stream being made up of separable sounds. This is why the teaching of phonics is such a key component of early language learning (see Section 1.2.1 where phonics is discussed more), and forms the baseline for reading ability.

Reading comprehension may not be just defined by understanding of meaning, as good readers have to be able to read well too. Reading fluency is an important aspect of reading comprehension (Godde et al., 2019), which is the ability to read text at speed but also ensure the meaning of what has been read is understood (Groen et al., 2019). Oral reading of text is an important aspect of learning to read for young children, partly because it is a common method for assessing children's reading progression at school, but also because it provides a cue to children themselves of how well they are reading (Kuhn et al., 2010). That is, children comprehend the text through their own oral reading. Reading appropriately aloud allows children to practice reading the text, and may expose difficulties with their reading ability. If reading for understanding is one of the primary goals of reading, then reading well is important to allow this to occur.

The growth and competency of reading comprehension is accompanied by a general increase in cognitive and language ability (Hoffman, 2009), but the focus here is on the SVR definition of reading comprehension. To reiterate, this is that reading comprehension is made up of two components, decoding and linguistic comprehension. Within the statement of the original framework Gough and Tunmer (1986) were using decoding and linguistic comprehension as ways to attempt to explain difficulties with both, such as poor decoding representing traits of dyslexia. Over time the framework has moved to be more of a representation of skills which correlate with variability in reading comprehension (Aouad and Savage, 2009).

One important aspect of reading comprehension which is important within this framework, is the comparison between linguistic and reading comprehension. On first thought, the word comprehension suggests both these aspects are simply the same concept of understanding, with only the presentation of information the difference (oral vs. written). This relationship between reading and linguistic comprehension (they use the term listening comprehension, which as discussed in Section 1.1.1 is sometimes used as alternative to linguistic comprehension) by Wolf et al. (2019). The study was designed to look whether listening comprehension and reading comprehension were the same general comprehension skill presented through different modalities, or different processes with different underlying modality-specific skills. Their initial discussion focuses on this, examples such as attention and memory skills may be more important for linguistic comprehension as this is spoken and needs to be recalled whereas written text can be revisited. Comparatively, reading fluency (the ability to read text with few errors and a natural speed) may be more important for reading comprehension (especially oral reading) as both meaning and decoding must be conducted in parallel to ensure the written word is interpreted correctly (Kim, 2015; Kim et al., 2011; Schild et al., 2014; Schrauben, 2010). These differences in skill set may complicate the exact comparison between linguistic and reading comprehension. This also leads to a reasonable critique of the SVR, that it does not satisfyingly differentiate these two skills, rather that they are two aspects of comprehension conducted in different modalities. It is beyond the scope of the current thesis to investigate the modality-specific differences between linguistic and reading comprehension, but it is an important aspect to note when the current project chooses to use a framework which seems to have issues in distinguishing between these two types of comprehension. I discuss this in Chapter 3, where I used the example mentioned in the previous section by Whalley (2017) of matching the linguistic and reading comprehension measures, to try and control for the differences between the tasks.

Reading comprehension can also be affected by factors outside the cognitive ability to just engage with text. For example, background knowledge, as discussed in the DIME model by Cromley and Azevedo (2007), will affect a reader's ability to comprehend a text. Take for example a research paper about biology, compared to a layperson a biology student has more information that will aid their understanding of the text, such as specific terminology that is being used. This allows an informed reader to integrate meaning with what they already know, likely making it easier for them to comprehend the meaning in the text. Thinking about developing readers, they likely do not have a lot of context for the text they are presented with. As such, reading comprehension begins difficult and becomes easier over time with more experience and knowledge of the world around them. Another non-cognitive aspect of reading comprehension is motivation, a multifaceted concept which explores a readers interest, involvement and perceived control over a text (Guthrie et al., 2006). Children's own sense of competence with reading will affect their ability and enjoyment, so finding a way to maintain this motivation is important (Wigfield et al., 2016). These are just two examples of non-cognitive aspects of reading comprehension that should be considered. In the context of the current study, interest was focused upon the cognitive aspects however.

The refrain of reading comprehension being the end goal of reading has been mentioned already a few times in this subsection. Reading comprehension is the culmination of children's journey of reading development, and will continue to improve throughout their early education. Within the context of the SVR, reading comprehension comes from the combined development of decoding and linguistic comprehension ability, which between the two definitions explored prior encompasses many language skills. This feeds into a critique of the SVR to be discussed in the following subsection, the framework allows for many aspects of oral language understanding to be brought under the umbrella of either decoding or linguistic comprehension, but this doesn't give a specific picture of what reading comprehension is. Reading comprehension is reading, but it is also so much more, and this is the conversation the SVR framework brings to the fore. Next

I explore critiques of the framework, and end discussing the lasting impact of the framework in the current state of reading research.

Summary

To summarise, the SVR is comprised of decoding and linguistic comprehension, which are two components which contribute to reading comprehension ability. Decoding aids this through the development of word reading ability, allowing for children to automatically view and read words without much cognitive effort. It may be that the predictive relationship weakens over time with the automaticity of decoding, and when reading ability moves from learning to read into reading to learn. The second component is linguistic comprehension, the ability to decipher and gain meaning from spoken language. This has been matched to reading comprehension, with the difference being the way the information is being presented. It is likely that children's ability to understand and garner meaning from speech will help them in comprehending the written word. The framework works as a snapshot of a child's reading ability at the time of measurement, it does not work as a developmental model. Next is to discuss why the framework has managed to maintain its relevance in the literature, and also the critiques of the framework from researchers.

Strengths and Limitations of the SVR

In a recent issue of the journal *Remedial and Special Education* in 2018 a special issue was commissioned to explore the SVR, with three papers exploring its usage. This was led by a paper by Hoover and Tunmer (2018), as already covered two of the major researchers behind the framework itself. They reflected on the history of the SVR, and thought about what it represents thirty years later, going so far as to label it a conceptual model, and discuss what it is not. They conclude though by highlighting they believe the "power" of the framework was both it made a complex phenomenon easier to understand, but that it was also falsifiable. They believe it still provides insight into the developing reader, but are willing to accept that it can represent reading at the broadest level, and that its continued relevance is of great comfort to them.

Why is this framework so useful for exploring reading development? The two components of decoding and linguistic comprehension tap into what it takes to become a successful reader, the ability to recognise words and to take meaning from them in context. In their seminal paper, Castles et al. (2018) reason that the framework is so useful because it shows that "decoding and linguistic comprehension are both necessary, and neither is sufficient alone". This is reflected by the diagram in Section 1.1.1, wherein reading difficulties are seen by deficits of one or both of these skills, leading to an overall problem with reading comprehension. The SVR itself does not deny the complexity of reading comprehension, but decoding and linguistic comprehension are two clear sets of individually complex processes (Catts et al., 2015), and provides the framework with which to explore a complex construct in a focused manner.

That said, not everyone agrees with this framework and its proliferation. Some see the SVR as a regression of sorts, ignoring the complex process of literacy and focusing it onto two components that are important, but lack detail that is important for a reader to flourish. One such critique is presented by Uppstad and Solheim (2011), who explored the limitations of the SVR as a framework of reading ability. One issue they raise is that equating linguistic comprehension and reading comprehension as similar processes, only the medium it is presented to the learner is different. Uppstad and Solheim highlight that equating these skills is an issue, because the difference in demands on the author between oral and written language is enough to place different demands on learners. Speech is an interactive experience, whether it is face-to-face with the aid of body language, or over a phone where you get the aid of prosody and contextual cues in the language spoken to aid comprehension, spoken language gives the listener information to help decode what is being said beyond simply understanding the individual words said. On the other hand, reading is for the most part a solitary experience, whether oral or silent reading it relies on the reader to both decode and make meaning of the text as they read. This goes beyond decoding and comprehension, to visual acuity and the context of the text being read. Whenever discussing reading there is an endless array of skills and components that can be discussed as to what makes a successful reader throughout childhood. The complexity of reading cannot be captured in a single framework, that is true, but it is hard to imagine any framework or model covering everything that occurs when an individual reads text (Castles et al., 2018).

James Hoffman reflected on the utility of the SVR in the Handbook of Research on Reading Comprehension (2009), and came to the conclusion that the SVR is inadequate as a theoretical framework for understanding reading, because it understates the complexity in developing reading comprehension (p.63). In an even stronger assertion of his position, Hoffman believes that it is wrong and misleading to utilise the SVR for teaching and understanding reading (p.64), in part because its only new contribution is as a simple view of a theory that can be easily understood (p.58). As a counter to Hoffman's conclusions, the important point to remember is that the SVR is a framework, it does not claim to be a theory to explain all of reading development. Rather, as Stuart et al. (2008) discuss, it may be that the SVR represents the skeleton for essential reading, and the flesh of reading is filled out through other models. As seen above, there is a consensus in the literature that the components of the SVR represent a common ground for the essential components of developing reading comprehension.

However, Hoffman's writing does highlight a weakness of the framework that I explore in the current project. That is ignoring prosody (the patterns of stress and intonation in spoken language) as a potential mediator between the decoding and comprehension aspects of reading ability (Hoffman, 2009, p.60), and how it could potentially fit within this framework to help us understand reading ability better. Prosody itself will be discussed more in Section 2.1, but for the moment it is important to understand that prosody refers to the elements of speech that are not segmental, rather prosody occurs at a syllabic or larger unit level, known as suprasegmental. In speech, it is reflected in intonation, stress and rhythm, with the individual controlling this. One example is distinguishing between a statement and a question, the intonation of how you say “Hello" and “Hello?" has different connotations. This makes prosody an important aspect of comprehending speech, but it is not mentioned in the SVR framework, and as such could be a missing aspect of the equation. Whalley (2017) reflected on this in their own thesis, but were not as critical as Hoffman was in their writing, reflecting rather that it is that prosodic knowledge is not commonly considered in the SVR relationship. Written languages do not contain the same prosodic information you get from spoken language (see Section 2.8.2 for a longer discussion of this and the idea of implicit prosody), in that the English written word does not explicitly imbue text with prosodic cues in how to read the text. This is a complicated matter, in that the placement of stress and intonation in the written word will provide clues to the meaning of how the written text should be read, but particularly for English there is not the benefit of accents for letters as in a language like Spanish. For successful reading comprehension the reader has to bring their knowledge of prosody to read with a natural expression, and to fully comprehend the text.

Summary

I chose the SVR as a framework to model the current study, as it provided a basis for exploring reading comprehension ability with two clearly labelled components. Furthermore, since the Rose Report in 2006, the SVR has affected the teaching of reading within schools, and thus I felt it was important to utilise this framework in our work with children for whom this framework was influencing their reading ability. Awareness of its limitations is important, befitting its name it is a simplistic view of reading. But, its utility is as a broad summation of what makes a successful reader has led to it retaining its place in both research and pedagogy.

The Cognitive Foundations Framework

Two of the original authors who were a major part of developing The Simple View of Reading, Dr William E. Tunmer and Dr Wesley A. Hoover, have recently proposed a new framework for helping to discuss reading difficulties, which is an extension of the original SVR framework. In Tunmer and Hoover (2019) they discuss their new framework, the Cognitive Foundations Framework (CFF), which is very similar in design to the SVR.

As seen in Figure 1.3 the CFF is a framework of learning to read, with the belief that language comprehension and word recognition are the major cognitive components underlying successful reading comprehension. The authors describe what they understand as reading comprehension to be the ability of an individual to take the literal and inferred meaning from a written text, which allows children to read for self-understanding. The authors acknowledge that this model solely deals with the cognitive foundations of reading, and not other factors which can affect its development, such as psychological factors (e.g. motivation) or societal factors (e.g. SES, access to resources), which are distal factors that will have ramifications on reading ability. The focus of

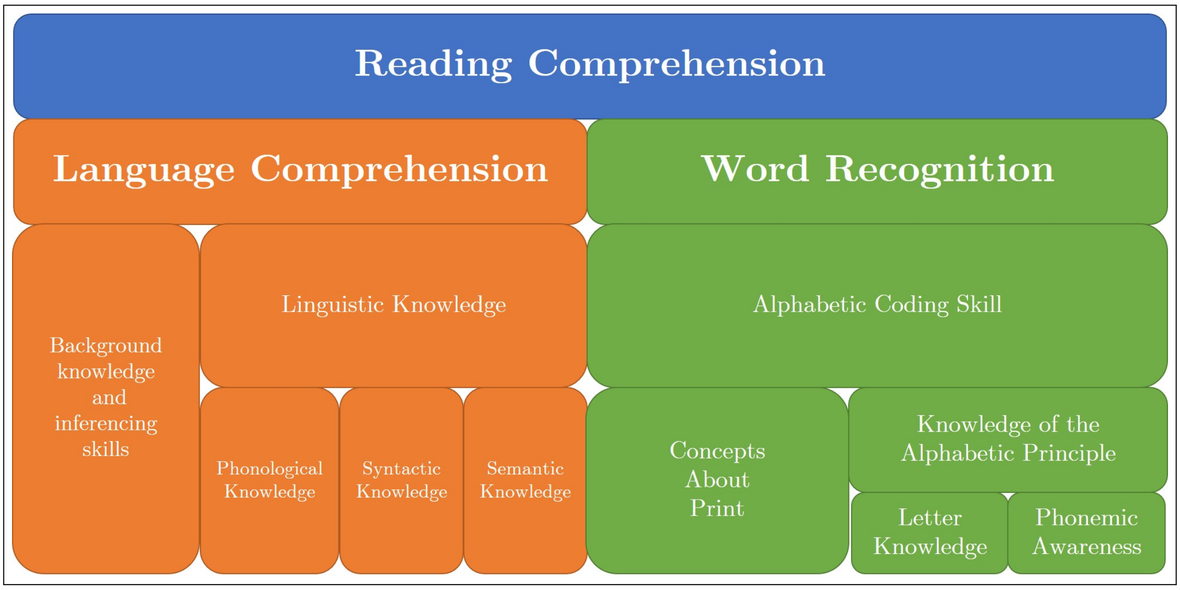

Figure 1.3 The Cognitive Foundations Framework: Adapted from Tunmer and Hoover (2019). The Language Comprehension and Word Recognition components align with the key components of Linguistic Comprehension and Decoding in the SVR, but extra detail regarding the cognitive background of these components is added to this model.

the current framework was to develop a framework for interventions, to help guide practitioners and other academics towards understanding the cognitive underpinnings of reading ability.

The framework itself is not meant to be interpreted as a hierarchy according to the authors, despite appearing that way. Rather, each aspect represents the overall skill, and the individual skills that make up this, for example to achieve word recognition one first should have some knowledge of the concept of print and the alphabetic principle before they achieve an automatised alphabetic coding skill. These lower-order skills may not be perfected, but some level of mastery has occurred, and a continued reciprocal development can occur as reading development moves on. This diagram is not clear about that though, and it is understandable from first appearances why this confusion could be found.

I am discussing the CFF because this new framework was described during the collection of data for this thesis, and builds on the SVR in some key ways to better define the components that contribute to reading comprehension. The current thesis has taken the SVR as a framework to decide on the choice of methods to explore the research questions posed. The CFF paper was released after the design of the project had been finalised, and it was interesting to read about this revision. The addition of the detail regarding the cognitive background of the components is helpful for deciding on the right measures to use for investigating the predictors of reading comprehension. Furthermore, the detail provided here supported my understanding of the SVR as a simplified framework by which to explore reading comprehension, and the CFF helps to provide a context for why this is so. It may be that this particular framework becomes a bigger part of future research if adopted by the research community. Though even Tunmer and Hoover acknowledge that the SVR has continued to be shown successful in exploring individual reading ability since its inception, and that this new intervention framework does not discount the old framework as guidance towards individual differences in reading comprehension across time.

As mentioned, for the current study I used the SVR as the framework for exploring children's reading comprehension ability. More details regarding the current study will be discussed in Section 3.1. For the moment it is pertinent to know that decoding and linguistic comprehension have been measured for this study, and are an important part of the research questions I have explored.

1.2. Oral Language and Reading

Twenty years ago Moats (1999) published a report for teachers discussing the teaching of reading, with a title to ensure readers took notice: *Teaching Reading IS Rocket Science*. The document was an attempt to compile information about reading research to help current and aspiring teachers on how to teach children to read. One does not need to go beyond the title to assume the thesis of the report, learning to read is difficult, as is teaching it. In the time since this report reading research has continued to develop an understanding of how reading develops, from difficult beginnings to a hopeful self-guided ending. But, in Moats' update to this her executive summary says that much of the reading research has not been included in teaching lessons or other professional development for pedagogy. This disconnect means that one of the most fundamental skills being taught to children still requires the research and pedagogy fields to communicate, to ensure children are given the best chance at becoming expert readers (Moats, 2020).

What this highlights however is that teaching children to read continues to be a complex issue, both in the academic research world and for those who teach children this skill. As seen by The Simple View of Reading framework, an important part of learning to read appears early in children's educational development, the oral language they use each day. In this section, I discuss at a high level the relationship between speech and reading, particularly with the English language, and why it is a particularly hard language to read for young children. Following this, I focus on a specific aspect of language ability, the use and understanding of phonology. I begin by talking about the importance of phonological awareness, and its prevalence in the literature investigating reading development with children. Then I discuss prosody, also known as suprasegmental phonology, in the following section. Prosody and its relationship to reading ability is one of the major aspects of the current study, and as such much of this section is dedicated to defining prosody, and exploring the work which has been done to understand the relationship between prosody and reading development. This section ends summarising the work explored up to this point, leading into a description of the current study.

1.2.1. What Do We Know About Early Reading Ability in English?

Reading Education in England and the influence of the Rose Review (2006)

In this project I focused on readers in England, and as such it is important to talk about the teaching of reading in the UK school system. There is no doubt that reading is an important skill for teachers to provide to their children, but knowing how to teach reading ability to every child is an ongoing difficulty. In her case study discussing the introduction of phonics to the school system, Stainthorp (2020) informs and reflects upon the changes in the teaching of literacy skills. The current curriculum was implemented in 2013 (Department for Education, 2013), and describes that the teaching of reading begins focused on two dimensions of ability, word reading and comprehension skills. This is familiar, having already discussed The Simple View of Reading, and the SVR was an influence on the current curriculum, that word reading (decoding) and language comprehension are two important aspects of early literacy ability. Specifically regarding the word reading aspect, one of the most well known developments in reading teaching in the UK was the introduction of phonics teaching, as was suggested by the Rose Review (Rose, 2006), the influential independent review of the teaching of early reading. The document supported the implementation of phonics, a method for teaching the alphabetic principle, in the early years of teaching.

The National Curriculum in England (Department for Education, 2013) specifically recommends that children engage with synthetic phonics programmes in the early stages of learning to read. This involves children learning about the individual phonemes of English, and learning how to sound out each phoneme, and then learning to manipulate (e.g. blending, deletion) the phonemes to create words. There remains to this day some controversy regarding phonics as a teaching method versus the old whole language model of teaching reading, a method of teaching reading where attention is paid to whole words rather than emphasising attention to the individual components. A summary can be read in a comprehensive paper by Castles et al. (2018). What is important to know for the current section is that phonics is the current system for teaching early word reading in schools in England, and that the method promotes a systematic approach to developing early reading ability. The Rose Review was informative towards to inclusion of systematic phonic programs into all schools for early reading. Though the improvement in literacy achievement was not immediate, 58% of children passed in 2012 when the phonics check was introduced, by 2017 those children who had gone through phonics training from the beginning of their education 81% were achieving the expected literacy level expected by the end of Key Stage 1 (Department for Education, 2018).

Related to the rise of systematic phonics teaching was the introduction of the Phonics Screen Check in 2011. Completed at the end of the first year of education, the check was to assess the progress of children in learning their phonics knowledge, by assessing their ability to read 40 words and nonwords. In her article, Stainthorp (2020) reflects upon the introduction of the Phonics Screen Check being met with controversy, partly due to a misunderstanding of the use of non-words in the test. It is understandable why, without appropriate context, one could question the validity of a test of reading ability which uses fake words to assess children. However, as Stainthorp clarifies, the assessment of phonics is not an assessment of reading, it is an assessment of how well children have taken on board the lessons related to phonics. The inclusion of non-words is to test children's ability to use their knowledge of individual phonemes to read words that do not exist, but contain combinations of phonemes that are readable. Stainthorp also adds that a truer test of phonics would be to only include non-words, as the children would have no possible contextual cues to help them decode the words in front of them. This would have been unwise due to the controversy of their inclusion already. The current study took inspiration from Stainthorp, ensuring I included a measure of nonword reading to analyse decoding. The interest in the phonics check for this project was to gain an understanding of the early development and assessment of decoding skills in the children we worked with, and how we might best probe this ability. This knowledge contributed to my choice of measures for the project, ensuring I included a measure of nonword reading based on Stainthorp's clarification as to the importance of this measure.

As well, the Rose Review recommended replacing the previously discussed searchlights model of reading, for The Simple View of Reading framework to guide teachers in the teaching of reading. This was discussed in the previous section, the searchlights model's weakness was that it did not provide a path for how a beginner reader becomes an expert reader, and so the SVR was brought in as a replacement.

There is more that could be discussed related to the teaching of reading, but this project focuses solely upon Key Stage 1 learners. It is important to highlight however that the teaching of reading is ever changing, due to the continued advancement and findings of reading researchers. Researchers should continue to ask themselves what they can do to help the continued understanding of reading ability in young children. Throughout the remainder of this section I will reflect upon the development of reading ability from oral language, and how the early years of understanding language come to influence later reading ability. Before I get to this, it is important to discuss the English language, and the particular difficulties and quirks it provides to developing readers.

Reading in English and the Alphabetic Principle

With the knowledge of the school curriculum coming to put literacy as a major focus for developing children, it is important to know how one comes to be a successful reader. It is a constant refrain echoed throughout the reading science literature that reading is not a natural skill, but one which requires time and effort to acquire. There are a multitude of avenues to explore within reading science, but this section focuses on English as a language for learning to read, comparing it to other languages which use the twenty-six letter alphabet, and why it is important for us to think about.

A major part of learning to read the English language is understanding that there is a relationship between the words written on the page and the language that is spoken. This is the alphabetic principle, the understanding that the visual elements of the English writing system (graphemes) represent the spoken sounds of the language (phonemes). This is one of the major starting points for English readers and may be the first time that children recognise that those symbols on the page are related to their speech. To acquire the alphabetic principle is dependent upon children understanding that words are composed of smaller sounds, such as phonemes, and an awareness of the alphabet itself (Petscher et al., 2020). This knowledge is particularly key for English, as being a morpho-phonemic language readers are particularly reliant upon their ability to segment words into phonemes and morphemes, to both decode and comprehend written language (Moats, 2010).

Explicit teaching of these principles seems to be key to the beginning of reading with young children, systematic phonics teaching an example already discussed in the previous section. The current project does not explore in depth the pedagogical side of teaching literacy, however from the science of reading perspective it is important to state that the transition from speaker to reader is not as easy as the children simply being aware of the alphabetic principle. As Snow (2020) states in her paper reflecting on the science of reading, there is not a clean transition from a child being a “good talker" to being a “good reader", more work needs to be done to make a “good reader" than one may assume. However, English is not the only language which depends on the twenty-six letter alphabet, why is the discussion of English so important?

Reading in English and Orthographical Differences

Alphabetic orthographies are all based around letter-sound relationships, but there is variation in how closely the orthography being discussed reflects the language’s phonology. In other words, languages differ in the relationship between the letters and sounds of their specific language. This thesis is focussing on children who speak British English, and it is important to note that English is a particularly difficult language for oral-to-reading learning. The spelling of a written word should reveal to the reader how to speak it aloud (Van Orden and Kloos, 2007, p.61), but this is not the case for languages universally, especially the English language. For example, the oral pronunciation of the word *through* is easily perceptible and learnt through speech with exposure, but when encountered initially in reading children are likely to be confused, as the written spelling does not match other words with a similar spelling and a more common pronunciation, such as *rough*. English is often cited as a great example of the inconsistent or deep alphabetic orthography, because of its “myriad of exceptions, irregularities, and context-dependent ambiguities" (Lyytinen, Shy, and Richardson, 2015, p.704) and because a phoneme can have multiple spellings, and a letter or combination of letters can have many pronunciations (Ziegler and Goswami, 2006).

A much-cited study by Seymour, Aro and Erskine (2003) compared orthographic acquisition in thirteen different European alphabetic orthographies, including English and Finnish. These two languages are notable in this study as they represent opposites in the complexity of learning to read in some respects. Finnish children learn to read relatively easily, as the written language has a simple syllabic structure and a shallow orthography (the written and spoken language overlap with each other consistently), meaning children's language proficiency prior to reading is highly predictive of their reading ability. English, on the other hand, is much more difficult, it has a complex syllabic structure and deep orthography. The English language is represented by 26 alphabetic letters but has over 40 phonemes to be represented from spoken language, so some sounds must be represented by complex graphemic units made up of more than a single letter (Lyytinen, Shu, and Richardson, 2015, p.714). A comparative example shows this well; the first phoneme in ship (ʃ) cannot be represented by a single letter and so is digraph, but this single grapheme will also represent the phoneme in other words such as in ship (ʃɪp) vs. shine (ʃɑɪn). This brief aside brings us back to Seymour's study, which found that children learning to read in English required, on average, an extra two years of explicit tuition compared to more transparent orthographies to read at a competent level. This is likely due to the extra rules children must learn to account for inconsistencies in the language, as well as learning the alphabetic system.

Despite the importance of spoken language for reading and its foundational development, to read goes beyond simply understanding spoken language (Whalley, 2017). When listening to speech the listener is given the advantage of the physical sound of speech, with the speaker's emotion and intonation clear, as well as the visual aid of the speaker's face to help comprehension. Though not all these cues are necessary for the comprehension of speech, the more of these cues you have to use the more likely an individual is to comprehend what has been said. In contrast, reading does not provide these clues to the reader; rather the reader must come to the written word and fill in the blanks themselves if not clearly marked in the text. Despite this hurdle, most children who begin learning to read become successful readers throughout early childhood.

In the following subsection the focus moves towards discussing the relationship between oral language and early reading development, with a particular focus on phonology.

1.2.2. The Relationship between Oral Language Ability and Reading: Focus on Phonological Skills

Before children can begin to develop their reading ability, they will have spent their youth developing their oral language ability. There is an intimate connection between reading and spoken language (Whalley, 2017) in which reading is an extension of the speech-language system (Ashby, 2016), such that oral language ability is often used as a predictive measure of later reading capability. Without this prior grounding with spoken language, the information provided by the written word is almost useless to children. Much research has been conducted exploring reading development, and it is clear literacy is very much rooted in oral language capability (Wade-Woolley, 2016, p.4).

To explore this relationship I focus on phonological aspects of reading ability. This is partly due to the large amount of work conducted investigating phonology's relationship to reading development, in particular segmental phonology which will be discussed in the following Section 1.2.3. The reason for this focus is that to explore each aspect of oral language with reading, such as syntax and morphology, is beyond the scope of this thesis. As mentioned, the impact of phonological knowledge is a well-researched and accepted aspect of early reading development, and it is important in understanding the design of the current project. Through this focused approach I can explore the aforementioned segmental phonology and its importance for understanding children's early reading skills, and then I move to discuss suprasegmental phonology, also known as prosody, in the following chapter. Children's understanding of prosody and its relationship to reading is the main focus of the current project, and as such this chapter will spend time defining what prosody is in the current project, what past literature can tell us about the relationship between prosody and reading ability, and finally the rationale behind the exploration of prosody and reading comprehension for the current study.

1.2.3. Segmental Phonology and Reading

An important aspect of children's developing speech system is their ability to interpret the speech stream of others in the world, learning to convert these seemingly meaningless sounds into units they can understand and comprehend from (e.g. words, syllables, phonemes). This is a simplified description of phonology, the unconscious awareness of how the sounds of spoken language are organised into patterns and systems. Through development and understanding of their languages phonology, children can understand the basic blocks of their language. Van Orden and Kloos (2005) posited that phonology acts as a mediator between the oral pronunciation stored in the brain, and the written word. In other words, the importance of phonology is that it establishes the link between the two. They represented this with a chain of events that could represent reading, starting from processing of the visual stimuli, through to decoding, creating the phonological representation and finally the motor system for pronunciation. Though this is a nice tidy chain of events, such a process is too neat to summarise as such. The complexity of reading acquisition goes beyond phonology alone, but at the same time it is critical for reading development (Carroll et al., 2011, p.1), as the written word functions as the visual representation of the language's sounds, and so to develop this understanding is important.

Children are aware of the phonology of their language from very early on in their language acquisition, maybe even before they are even born. A study by Mahmoudzadeh et al. (2013) worked with babies who were born prematurely during the third trimester, and measured their brain activity with near-infrared spectroscopy, a non-invasive method of measuring brain activity. They wanted to investigate whether these children were able to distinguish two syllables ("ba" and "ga") that differed only in the phonological contrast between them. What they found was that this phonetic contrast did produce different responses in the brain's language areas, similar to adult linguistic computation, suggesting these babies had some learnt or instinctual processing of phonological information. Furthermore, Mahmoudzadeh et al. found the change in human voices (phonemes read by a male or female voice) uniquely activated differing areas of the left frontal region of the brain.

Recognising the importance of oral language ability to the development of reading, researchers have spent much of their time exploring how understanding of phonology relates to early literacy. The most well-researched aspect is that of segmental phonology, referring to the separable sound segments of speech, which are commonly defined in the reading development literature as; words (e.g. *cat*), onset and rime (e.g. *c* and *at*), and phonemes (the separable units of sound in a language e.g. /k/, /a/, and /t/ are the three phonemes that make the word *cat*). This is an important aspect of alphabetic languages in general, the knowledge that they are made up of individual segments that can be changed and rearranged to make new words. Taking the example of ’cat’ above, we can then proceed to change each phoneme individually (’bat’, ’cot’, ’cob’), in turn changing the meaning of the word. Knowledge of this is powerful, it is incredible that the language is so malleable with just 26 letters. English has a disparity between the number of letters we have and the phonemes that can be used, but overall it provides a building block towards understanding language.

Phonological Awareness and Reading

The most widely researched aspect of segmental phonology in the last few decades in early reading development is phonological awareness. This is defined as the awareness of the phonological structure of language (Marinus and Castles, 2015, pp.663) and the ability to understand the relationship between the orthography and phonology as separate entities in reading (Lyytinen et al., 2015, pp.713). Phonological awareness is a well-reported and evidenced predictor of early reading ability in young children (Bradley and Bryant, 1983; Castles and Coltheart, 2004; Cunningham and Carroll, 2015; Hulme and Snowling, 2013; Melby-Lervåg et al., 2012). There are many studies providing evidence that children who score better on measures of phonological awareness also score better on tests of early reading ability, most commonly with word and non-word reading tasks requiring the ability to decode written text (Hulme and Snowling, 2013; Melby-Lervåg et al., 2012). It has been described as a foundational literacy skill for understanding the alphabetic language (Van Wingerden et al., 2018) and one of the “most potent ingredients" in learning to read and spell in English (Wade-Woolley, 2016).

The fundamental challenge young readers face is converting the graphemic representation of their language into a phonologically based representation, which allows them to access their knowledge of words (Shapiro et al., 2013). Thus, children require a strategy that allows them to quickly decode the letters and words on the page to allow for fluent and automatic reading ability. Konza (2014) describe this process succinctly; children go from seeing the word cow as an animal, to a word, a word that has syllables, and then to separate the single sounds into their phonemes. Successful phonological awareness teaches children that linguistic units are made up of various sizes (Kilpatrick, 2012). With this knowledge of linguistic ‘grain sizes’ (Ziegler and Goswami, 2005) children are then able to manipulate and decode the orthographic representations of words they encounter and become ever better at decoding the phonological representation of written words, eventually becoming able to match letters to sounds and independently attempt to sound out new or unusual words (Whalley, 2017). Furthermore, children’s competence with phonological awareness develops typically along this grain size gradient i.e. children often develop rhyme awareness before phoneme awareness (Ziegler and Goswami, 2005).

Phonological awareness is but one component of a larger phonological processing system, which includes phonological recoding (recoding written symbols into a sound-based representation) and phonetic recoding (working memory system), which are both important to reading development (Wagner and Torgesen, 1987). However, it is phonological awareness which has been most well-represented in the literature as contributing to reading development from an early age. Melby-Lervåg et al. (2012) proposed that this is likely because this skill requires children to explicitly reflect upon and manipulate the speech sounds they encounter, which in turn aids children’s phonological development in the process. Phonological awareness appears to have a critical contribution to make to reading development and has been widely regarded as a useful skill to focus upon in the discussion of reading ability (Whalley, 2017).

Despite the ubiquity and density of research investigating phonological awareness, expertise at this skill cannot account for all the variance in reading development ability (Goodman et al., 2010). This alternate research pathway is the focus of the current thesis and the following chapter, prosody.

Chapter 2. Prosody and Reading

The focus upon segmental phonology discussed in the previous section is important, as the majority of research exploring the developmental relationship between oral language and reading has focused upon this. However, the current project focuses on another aspect of phonology which is essential for successful oral communication, but has often been neglected in work investigating the science of reading (Gross et al., 2014).

This is prosody, also known as suprasegmental phonology, the ability to imbue text with extra meaning and context above simply just the word itself. In a more practical short explanation, Veenendaal et al. (2014) define prosody as not *what* is said (i.e. the words themselves) but *how* it is said. They expand on this that prosody is not just about conveying meaning through speech, but that it makes speech much more comprehensible and natural for listeners, in turn making it easier for meaning to be extracted. But even this definition does not begin to encapsulate what prosody actually is. The first section of this chapter is an attempt to define what prosody *is*, and inform the reader how I chose to define prosody through the rest of this thesis. This is related to how I chose to measure prosody, and the information you want to explore with this output. Following this, I discuss three major components of prosody, that will be common when I discuss the research that has been completed in this area; stress, intonation and timing. This then brings us neatly to exploring the relationship between prosody and reading development, summarising work which has explored the relationship between prosody and reading ability, with a particular focus on word reading and reading comprehension. Discussion is also had of some of the multi-component measures used to analyse prosody, and how this is an aspect of measuring prosody which is under-used in the field. I will then discuss research which have tried to model the prosody and reading relationship, and discuss where this work could go in the future. I then highlight some other aspects of prosody research, namely prosodic reading and implicit prosody, that were not included due to the scope of the project, but are important to acknowledge and discuss as they represent fruitful areas for further research with children, as well as the wide range of research investigating prosody-reading relationships. Finally, I discuss how prosody will be explored in the current project, based upon the background reading conducted prior to designing the study.

2.1. What is Prosody?

Prosody is essential for communication, it provides us with the subtleties of meaning, the ability to represent unique quirks of language through a slight difference of prosodic form. Researchers often like to quote from Wennerstrom (2001), with her description of prosody as the "music of speech", an evocative and not untrue statement, as without prosody speech would be alike to the early monotonous speech-to-text parsers from years gone. Carlson (2009) furthers this, adding that prosody is "the tune and rhythm of speech" (p.1188), which is somewhat accurate as will be seen.

Prosody is an aspect of suprasegmental phonology, features which extend over more than one speech sound (e.g. syllable, word, phrase) and contribute to meaning (Miller and Schwanenflugel, 2006). The prefix ‘supra’ means ‘over and above’, which is essentially what prosody’s relationship to the segmental phonology previously discussed. Prosody works above the sounds of segmental phonology, to bring a richness and texture to speech. The descriptions up to this point have been very general, regarding what prosody actually is. This is because though the functions of prosody are essential to successful communications (Godde et al., 2019), they are complicated to explain with one definition.

At its most simplistic, prosody is the act of speaking with feeling (Paige et al., 2017), and this is the viewpoint most commonly associated with prosody. Gross et al. (2014) demonstrates this with the example of "Brian bought a book"; "Brian bought a book" is a statement, "Brian bought a book?" is a question, "Brian bought a book!", an exclamation, and "Brian *bought* a book" could be regarded as sarcasm. Prosody is not only allowing the speaker to express meaning and their thought process, but it provides the listener with extra information from the speech stream to understand the emotion of the speaker and allow them to respond appropriately. In short, prosody aids the segmental aspects of the speech stream by giving the speaker and listener extra choices in how best to represent the words they want to speak. Though this definition is useful to partially explain what prosody is, it does not reflect its full complexity. Prosody is a multi-faceted term (Lochrin et al., 2015) that refers to an aspect of oral language that means something different to different researchers, and this section will attempt to explore definitions of prosody used within the field of developmental reading research.

Gutiérrez-Palma et al. (2016) recognised that the definition of prosody is multi-faceted depending on the context of the research, and put forward two possible approaches to discussing prosody. The first is to look at how prosody supports the phonological subsystem in representing speech into hierarchical units, assisting the syntactic structure of speech e.g. syllables, phonological words and utterances. The second is to define prosody as the realisation of prosodic features, such as stress, intonation and timing. This second definition provided by Gutiérrez-Palma et al. is what was explored in this project, the concern being with the linguistic implications of prosody.

Three Components of Linguistic Prosody

Following on from the definition discussion above, reflecting on the second definition regarding prosodic features that contribute to understanding speech. The first feature to discuss is stress. Stress is typically marked in language by one syllable being more prominent than the remaining syllables (Becker et al., 2018). Stressed syllables are identified as being louder, longer and of higher pitch than unstressed syllables. Goodman et al. (2010) highlight the importance of stress sensitivity to English children, as stress patterns are central to oral language in the English language to differentiate meaning. A simple example is with the word ‘record’, with stress on the first syllable (REcord) the word is regarded as a noun ("Play that record please"), whereas when stress is put onto the second syllable (reCORD) suddenly the word is used as a verb in context ("Let’s record a song"). Children’s ability to successfully apply stress to the words they speak, and interpret stress contrasts from another speaker leads to successful communication with others. Indeed, it may be that stress is the most important prosodic component in the English language (Holliman et al., 2017b), it is certainly the aspect of linguistic prosody most well represented in the literature.

The second feature to discuss is intonation. This can be identified by the contour of the rise and fall of pitch in speech (Kim and Petscher, 2016), and is what may be referred to as the musical aspect of speech (Kent, 2013). Intonation aids oral language use as an aid for indicating meaning or intention through changes in pitch through the utterance, as well as telling the listener when a phrase or sentence is about to end. An example to display this is the contrast between statements and questions; the difference between "Hello", ending with a falling pitch, and "Hello?", ending with a rising pitch. The pitch change the speaker chose to use in context will affect the meaning the listener receives. Intonational cues such as the above example provide the speaker with the ability to add extra meaning and information to the speech stream to aid comprehension. The absence of intonation in speech is detrimental for native English speakers (Gross et al., 2014), as it removes major clues and context to speech and makes it difficult to gauge an appropriate response.

The third feature of prosody to discuss is timing. This aspect may be the least obvious aspect to identify, as it is a subtler aspect of prosody that can have an impact on meaning. One example from Holliman et al. (2014a) is discriminating between compound nouns (e.g. ice-cream) and noun phrases (e.g. ice, cream). Such differences in the timing between these words leads to very different meanings when put in the context of a sentence. Perception of timing properties, such as pause duration, aids to pick apart and comprehend the meaning of sentences.

As I continue to discuss prosody throughout this chapter these aspects will be highlighted in various measures and studies discussed. However these three components are a simplified set of prosodic features. The studies discussed below focuses on behavioural studies of prosody, that is research in the field of psychology that attempts to capture prosody in tasks with correct/incorrect answers, operationalising prosody in such a way that does not involve analysing speech for example.

2.2. Acquisition of Prosody

Prosody is an essential linguistic function for communication, and as such sensitivity to prosodic nuances from an early age is crucial for speech development. In their first year of life, infants pass throw a multitude of milestones towards mastering their native language (Quam and Swingley, 2014). Even from infancy babies rely on prosody to understand speech and learn to speak, with particular attention paid to the mother's prosody in early life. Even when newborns are three to four days of age they have been shown to show preference to their native prosodic pattern compared to non-native languages (Christopher et al., 2001).

Prosody is interwoven into language development, and to show its importance there is a hypothesis known as the prosodic bootstrapping hypothesis. This hypothesis describes that newborns, infants and children use the rhythm/timing of their language to identify words and untangle the speech stream, and as such is one of the initial building blocks for learning language (Whalley, 2017). The explanation this hypothesis holds is that newborns and infants can 'bootstrap' prosody and the cues it provides for understanding the speech stream, that is the sensitivity to stress and pitch helps them to attune to their native language and facilitates the segmenting of the speech stream, and thus a developing vocabulary. This is because prosody emphasises the regularity and salience of language patterns, which infants are said to latch onto to aid them attuning to their language's structure (Whalley, 2017). Furthermore, the hypothesis assumes the language learning does not have any prior linguistic knowledge (Soderstrom et al., 2003), with infants developing a nonlexical method of interpreting their native language. These prosodic cues help infants to understand the syntactic structure of their language, and though they don't develop a full understanding of their native language's syntax prosody can provide the building blocks. It is important to note that the hypothesis does not denote prosody as the only important aspect of language learning, only that it provides the early nonlexical context to give children the initial ability to bootstrap language from the what is at the time a meaningless set of sounds.

A well-researched aspect of prosody in infancy is that of lexical stress development. An influential study by Jusczyk et al. (1999) worked with English-speaking infants and found that 7.5 month old children were sensitive to words with a strong-weak stress pattern, and by 10.5 months sensitive to weak-strong stress patterns. A more recent study by Curtin (2010) were able to recognise novel words based on their stress information, and could detect when stress had moved to an adjacent syllable (e.g. BEdoka becoming beDOka). At around two-to-three years old developing children can use duration, intensity, and fundamental frequency to contrast strong and weak syllables, though not with an adult-like production (Arciuli and Ballard, 2017). Sensitivity to stress is a particularly helpful aspect of language for infants to require, as they learn early that 90% are stressed on the initial syllable which helps with understanding the speech stream (Cutler and Carter, 1987). This early sensitivity to stress is essential in English for infants to develop understanding and comprehension of the speech stream.

Quam and Swingley (2014) discuss another aspect of prosody in their literature, that of the use of pitch cues in speech. Termed intonation in the previous section, this nonlexical aspect of speech is not be so relevant for word recognition early in development (Singh et al., 2014). Some examples that Quam and Swingley (2014) include that despite children talking about emotions like happy or sad very early, children don't seem to exploit pitch cues in their speech as adults do until the age of four (Quam and Swingley, 2012). This delay was particularly for English speaking children, and it is possible that because English is a stress-based rather than tone-based language this could be why the use of pitch contours is a later addition (Quam and Swingley, 2014). But, despite children not having an adult-like understanding of pitch cues in English, there is evidence that pitch does help infants in being sensitive to phrase boundaries at two months (Quam and Swingley, 2012), and that by six months children are using pitch contours to parse utterances into clauses (Seidl, 2007).

There is much more to explore in the acquisition of prosody, but the acquisition is not the focus of the current project. The important context to take is that even from immediately after birth prosody is aiding language development, and as such is important to explore how this aids language (and reading development) through children's early life.

Summary

Prosody is important for infants to develop understanding and use of from an early age, and from the research described this learning begins from the moment a child is born. This could be word recognition through sensitivity to lexical stress in their parent's voice, using pitch cues to help with interpreting the speech stream, the work to process prosodic information begins immediately. With this said, infants still need to develop to become adult-like in their understanding and usage, and that this adult-like development emerges with fewer pauses and interruptions in speech and the ability to use a more varied pitch range to express meaning and intent (Valle et al., 2013).

2.3. Modelling the Prosody and Reading Relationship

It is important to model the relationship between prosody and reading ability, to try and operationalise how prosody contributes to reading development, whether directly or mediated indirectly with other language skills. A conceptual model of literacy development (see Figure 2.1) was created by Wood et al. (2009) which focused upon incorporating prosody into the model alongside three other oral language mechanisms, namely: vocabulary, phonological awareness (both phoneme and rhyme), and morphological awareness. The authors indicated that prosody was important to include as a new variable as the segmental phonological abilities, though useful for single syllabic word decoding, were not satisfactory in explaining how polysyllabic words are decoded (Holliman et al. (2010b)). Prosody is not an afterthought in oral language, it potentially aids understanding of reading development in an important manner.

Looking at the paths more closely, each has a theoretical reason for existing. The first pathway, through vocabulary, is related to the idea of the periodicity bias (Cutler and Mehler, 1993), the idea that children are able from an early age to understand and use rhythmic properties of speech in their environment. With this understanding, they can bootstrap this knowledge in their word recognition ability, which in turn provides support for vocabulary development, and later phonological awareness towards reading success. We see this path in the model from prosody → vocabulary → phoneme → reading. As established, phonological awareness is a skill that has a long history in predicting later reading ability. The second pathway is as above, except that prosody contributes directly to phonological awareness, and through this to reading ability. The authors reflect that sensitivity to speech prosody may aid the identification of phonemes in words. An example of this is stress sensitivity, as discussed previously the differentiation between stressed and unstressed syllables with prosodic understanding can help with the speech stream, and it stands to reason this ability to differentiate stress in phonemes would benefit phonological awareness also. The third pathway is also a segmental phonology skill, a direct pathway with rhyme awareness, that is the ability to identify onset-rime boundaries, and how the peak of loudness in a syllable may correspond to a vowel location. The fourth pathway is a direct pathway to reading through morphological awareness, another potential predictor of reading ability.

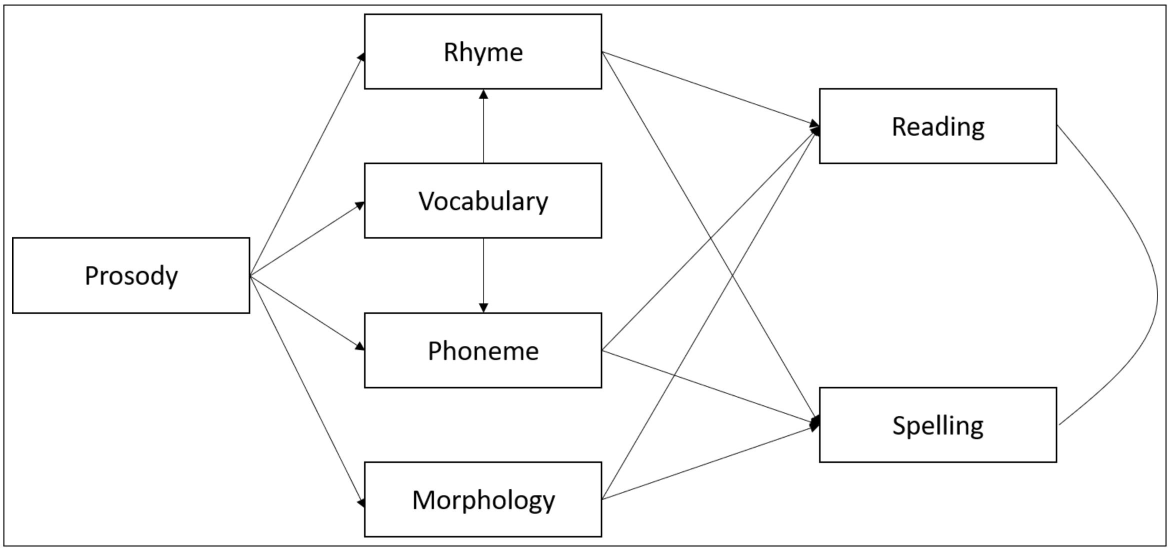

Figure 2.1 Wood et al. (2009) Conceptual Path Diagram

The first test of this model to discuss was by Holliman et al. (2014a). Two of the original authors of the model, Andrew Holliman and Clare Wood, revisited the model based on a sample of 75 five-to-seven-year-old children the authors recruited to test the relationship between these oral language mechanisms and their ability to explain reading and spelling. To measure prosody, the authors also included a more holistic measure of prosodic sensitivity, the Dina the Diver task. The measures of stress, intonation and timing were totalled in a variable simply called ‘Prosody’ in the final model.

The authors found some support for the model, with both the rhyme measure of phonological awareness and morphological awareness strong mediators between prosody and word reading ability. Otherwise, through their use of path analysis, they found that the model failed to fit their data. This is when the authors revisited the original model, and updated with four new additional pathways that made the model more complex, but perhaps more representative of the relationship of prosody and these other components of speech and reading. These four pathways were; from the rhyme measure to the phoneme measure, the rhyme measure to the morphology measure, the vocabulary measure to the morphology measure, and finally the phoneme measure to the morphology measure. The rationale stemmed from the feeling that the original model, though focused on the mediator effects of vocabulary, morphological awareness and phonological awareness was still too simple a representation of these relationships, which led to the addition of the aforementioned extra pathways. This new model was an excellent fit with their data, and the included variables accounted for 50% of the variance in word reading. Looking closer at the pathways however, the mediator pathways from prosody to phoneme and prosody to morphology were insignificant, with only the prosody to rhyme (r = .40) and prosody to vocabulary (r = .52) pathways significant in this case. The model suggested that prosody's relationship to literacy was complex, and occurred through a complex pattern of relationships with the paths discussed.

There were some limitations to this redesigned model that could be altered. The model did not include any test of the direct relationship between prosody and word reading ability, a relationship which has been tested in the literature (Calet et al., 2015) as discussed in Subsection 2.4, but still needs more work exploring. Furthermore, an improvement to the model may have been to investigate bidirectional relationships between the variables. This is a complex issue to solve, creating a model to successfully view and analyse the contribution of these variables between each other, and it may be that such a model will take many iterations to figure out.

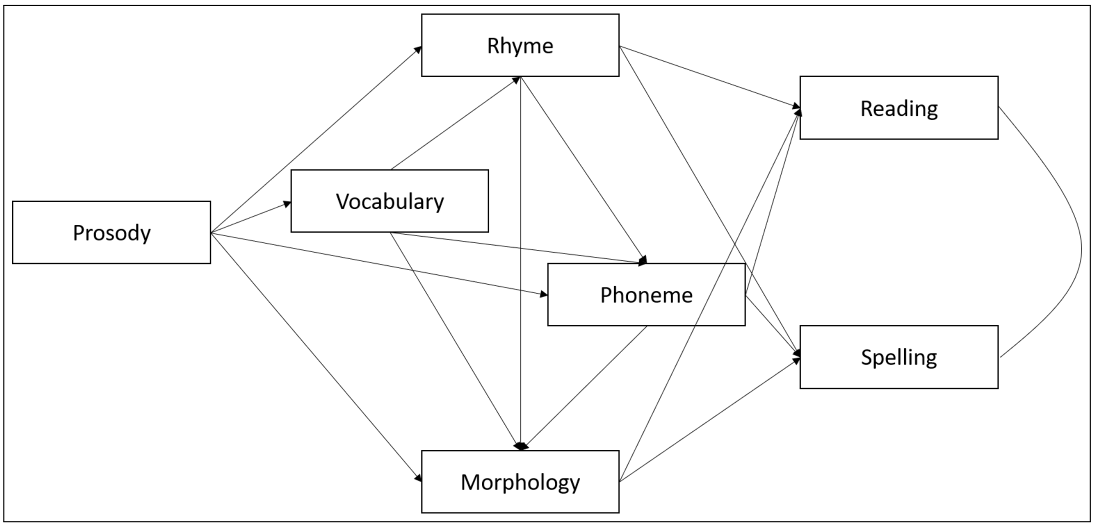

Figure 2.2 Holliman et al. (2014) Conceptual Path Diagram: An update to Figure 2.1, with four additional pathways. The first is from rhyme to phoneme. The remaining pathways are from rhyme, vocabulary and phoneme to morphology.

A recent paper by Critten et al. (2021) also looked into the Wood et al. (2009) model of prosody and reading ability, focusing on prosody's relationship with reading as an indirect variable mediated by vocabulary, phonological awareness, and morphological awareness. They used the modified model designed by Holliman et al. (2014a) to explore their results. The cohort was made up of 101 pre-readers with a mean age of four years and eight months, who were followed up one year later in their first year of school with a mean age of five years and nine months. In the first instance, researchers collected data about children's word reading ability, phonological awareness, vocabulary, and morphology. They used the Brenda's Animal Park measure of prosody designed by Holliman et al. (2017a). A year later, they collected data about children's word reading and spelling ability. The initial correlations found prosodic sensitivity to be positively correlated with word reading one year later (r = .26).

The relationships between prosody and word reading was then assessed with path analysis, based upon the model in Figure 2.2. The model was created by inputting the variables measured at time one (Prosody, Rhyme, Vocabulary, Morphology and Phoneme) and at time two (reading and spelling), with the path diagram drawn to see which of the pathways prosody had a significant contribution to the prediction of later reading and spelling ability. They found the model to be an adequate fit to the data they had collected, and that the variables measured at time 1 accounted for 25.4% of the variance in the word reading scores. What they demonstrated, was the model in Figure 2.2 provided an adequate prediction of how children's pre-reading prosodic sensitivity could predict word reading ability through its inter-relation with the other skills which have been seen to predict literacy. An important aspect in the design of this project was the use of longitudinal measurement with pre-literacy children, The authors highlight however, that a notable difference in the path analysis was that phoneme awareness alone directly predicted word reading ability, such that each of the other measures passed through this as the mediator towards word reading, including prosody. They propose this may be due to differences in the age group tested, being that they were younger and pre-readers, compared to the five to seven years of Holliman et al. (2014a). This only furthers highlights the need to explore prosody, and indeed reading development over an extended time span to explore how this changes over time.

Summary

I chose to summarise and discuss these three research papers and their modelling attempts, as firstly they are interesting pieces of research that contribute to the knowledge of the prosody and reading relationship. But further to this, they further solidify the difficulty inherent in trying to model the relationship between aspects of speech and reading ability. Add to this, a focus on prosody itself, which as discussed is a difficult aspect of language to operationalise and measure, seen by the Dina the Diver measure and its attempt as an inclusion measure of the aforementioned three components of prosody which has low internal reliability. These models may still be useful however, with cross-sectional and longitudinal studies designed with more participants they may pave the way to further understanding of this relationship. Indeed, over three studies there are small developments and realisations of how to tweak the model, or at least see how the model shifts dependent on the age group and measures used. It is hoped further research will utilise this model, and attempt to explore the relationship between prosody and reading.

2.4. Prosody and the Relationship to Developmental Reading

With a baseline set for the definition of prosody in the current study, the next step is to discuss the research that has explored the relationship between prosody and reading ability. The relationship between prosody and reading is not obvious at first, but when you begin to consider how individuals garner meaning from the text it becomes clear that prosody is important for reading. Prosody's relationship to reading is defined by Ashby (2016, p.68) as "the imposition of the intonation, stress, and rhythm of speech onto written text". What the field currently struggles with is understanding how the mechanisms of prosody are influencing reading ability, determining whether this effect is direct and/or indirect via other mediating variables. This has been difficult, as unlike phonological awareness prosody is not simple to measure. Prosodic features occur over several phonemes (Männel et al., 2017) and are multifaceted in their functions (Lochrin et al., 2015), it is no wonder researchers find it difficult to isolate prosody in developmental studies.

This difficulty is reflected when looking into where and when prosody has been mentioned in the literature. In 'The Science of Reading: A Handbook' by Snowling and Hulme (2005) the term prosody is only mentioned five times in total, and most of this discussion is about its omission from the field of reading research (Gross et al., 2014). Treiman and Kessler (2005) have described the omission of prosody in the discussion of reading development as unsurprising, not because it isn't important, but because writing systems generally ignore marking prosody in their orthographic representation of the language. Taking English as the example, there are no distinct markings to distinguish length, tone, stress and intonation when reading. Punctuation can serve as a guide to a point, but it isn't a totally reliable guide in comprehending what is being communicated in the text. This means that there are very few studies about how prosody occurs in early literacy (Beattie and Manis, 2013), but there are a growing number of studies investigating this relationship, attempting to pry apart the relationship.

I note now that this section focuses upon studies which investigated the relationship between prosody and reading comprehension in English. This means I am ignoring a large expanse of knowledge in other languages, including Dutch (Groen et al., 2019; Veenendaal et al., 2014, 2016), Spanish (Calet et al., 2019; Defior et al., 2012; Gutiérrez-Palma et al., 2016, 2009; Gutiérrez-Palma and Reyes, 2007), and other languages I did not cover in the literature search, such as tone-languages where the use of prosody is of major importance in meaning-making.

2.5. Prosody and Word Reading

Much of the research has focused upon word reading as the measure of reading ability. This is likely due in part to word reading tasks being simple to implement in research with children, and that they represent an early aspect of reading ability, that is the decoding of words. Furthermore, researchers can also use non-word reading tasks, wherein children read words that are decodable but do not exist. These tasks rely on children's internalised knowledge of the phonemes in their language, and the ability to blend them into a comprehensible spoken form. Prosody and reading comprehension's relationship will be explored in the following subsection, as there is more to discuss and explore with this.

2.5.1. Prosody, Stress Understanding and Word Reading

Word reading is a common measure used for representing reading ability as discussed, but when authors have chosen tasks to represent prosody, in our reading this has mainly focused on stress. I have already discussed stress perception as an important aspect of defining prosody, and how this awareness aids reading ability. It makes some sense that researchers would choose this, as assignment of stress in words is an important aspect of decoding them for pronunciation and comprehension.

The first of these studies was by Holliman et al. (2008), wherein the authors explored if five and six year old children's sensitivity to speech rhythm (prosody, as measured by stress manipulation ability) explained unique variance in reading ability. This speech rhythm task was the Mispronunciations Task, used originally in another study by Wood (2006). In this task children were shown an illustration of a house and introduced to a toy character who would guide them through the task. Children listened to a tape which randomly played a track with a recording of voice naming one of the objects in the house (i.e. sofa). All the objects used in the mispronunciation condition were two-syllable words, with stress on the first syllable, and a vowel reduction on the second. Once this baseline score was taken, children would return to the house with different toy characters, who would mispronounce the objects in the house in different ways. There were four possible mispronunciations possible via the following manipulations: 1) the location of primary lexical stress (first or last syllable), 2) the number of vowels changed (one or two), 3) the number of reduced vowels in the word (one or none).

Children heard this mispronounced word and were then asked to select the object in the house they thought represented what they had heard. Word reading was measured via two measures, a word reading task where children read as many single words as possible out loud, and a non-word reading task where children had to decode words that could be read aloud but were not real. In their regression model, Holliman et al. represented reading as a composite of the word and non-word reading tasks together called reading attainment. After controlling for age, vocabulary and the phonological awareness measure, prosody showed an independent contribution to the reading attainment measure ($\beta = 0.3$, $t(38) = 2.264$, $p = 0.029$), suggesting stress sensitivity has an association with reading ability independent of its relationship with segmental phonological awareness.

In a follow-up study, Holliman et al. (2010a) again investigated speech rhythm and its relationship with reading skills. Again, Holliman et al. used the Mispronunciation Task, but chose to revise it to be simpler for younger children and included more distractor items that began with the same letter sound and phoneme. In this task, children were not presented with a house with a large selection of items to choose from, rather the researchers presented children with four pictures of two syllable words, which each started with the same letter and sound (i.e. singer, skateboard, swordfish, and seagull). Then, by listening to the pre-recorded speech of a word, they would select which they thought they heard. The manipulation came by reversing the metrical stress of each word, so the first vowel became reduced and the second vowel fully articulated. Children were told they would be hearing words which would not be said properly, and they should point to the picture they felt best represented what they heard. They found that this new task correlated very significantly with their measure of word reading ($r = 0.63$, $p < 0.001$), and with a regression analysis provided a small amount of significant variance if entered at step 6 ($r^2 = 0.026$, $p < 0.01$) or step 7 ($r^2 = 0.021$, $p < 0.05$) after all the other control measures of age, vocabulary and phonological awareness. More research conducted by Holliman and colleagues will be discussed in the coming sections, as their more recent studies have focused on many other aspects of the prosody-reading relationship, and will span this whole section of the thesis.

Next is a project by Goodman et al. (2010) who also investigated linguistic stress sensitivity, and its relationship to both phonological awareness and reading development with Canadian pre-school children. The projects studied so far have worked with children who have had some taught reading experience, even if they are aged five or six in their first years of school. How readers come from having no knowledge of reading to being readers is important, and so research into this initial stage of learning is important. A major goal of Goodman et al. was to see if linguistic stress sensitivity would contribute unique variance beyond the phonological awareness task's predictive ability. 45 children from a single senior kindergarten in Ontario were worked with to complete the tasks. Lexical stress sensitivity was measured by the aforementioned Mispronunciation Task, namely the original version by Wood (2006). Metrical stress sensitivity was measured by the Compound Nouns task, which was an assessment of how well children could distinguish between compound nouns (e.g. ice-cream) and noun phrases (e.g. ice, cream). The control measures were a nonverbal intelligence task, a measure of vocabulary, and three individual measures of phonological awareness (phoneme deletion, phoneme blending, sound matching). Reading was again represented by a word reading task, but since these were younger children the task began with letter-knowledge (i.e. naming individual letters) and developed to words, as this was deemed appropriate for the age group. To assess the independent contribution of prosody, hierarchical multiple regression models were created. The first model to predict reading ability included verbal and nonverbal intelligence at step one, with metrical stress added at step two adding a small but insignificant contribution to reading ability ($r^2 = .03$,

$p < .05$), whilst lexical stress at step three was a positive significant predictor in this model (r^2 = .15, $p < .01$). This was then followed up with a similar hierarchical regression model, but this time including the phonological awareness measure as a separate variable. With phonological awareness as a predictor at stage three of the model, lexical stress was no longer a significant explanatory predictor variable to the measure of word reading (r^2 = .001, $p > .05$). This highlights the previously discussed important contribution of phonological awareness to early reading ability, it may have been that children's sensitivity to phonemes was more important than their stress awareness at this point in their reading development. Holliman's studies with slightly older children may reflect children who have started to internalise their phonological knowledge, and this stress awareness becomes more fundamental for their level of reading ability. It may also be related to the complexity of the words being read, where stress awareness is more important for more difficult multisyllabic words (an aspect of word reading to be discussed).

With the three studies discussed above, prosody has been represented by a measure of stress awareness accounted for significant variance in word reading ability, though with the presence of phonological awareness for the pre-reader participants this was not the case. Each has used the Mispronunciation Task as a representation of stress sensitivity, but there are more measures of stress awareness in the literature. Arciuli (2017) investigated this relationship further, but split stress awareness into two categories, dominant and non-dominant stress awareness. The participants in the study were 192 children aged between 5 and 12 years of age, recruited in Sydney, Australia. Arciuli developed a new measure of stress sensitivity, to investigate if differing patterns of stress and children's ability to understand them in turn led to successful reading ability. This measure is known as the Aliens Talking Underwater measure, which consisted of a series of two-alternative forced choice trials in which children saw a picture on the computer screen (e.g. "Coconut"), followed by an 'alien' saying two versions of the picture name, one with correct lexical stress (e.g. "COconut"), and the other with the stress across the first two syllables reversed to be incorrect (e.g. "coCOnut"). A low-pass filter was applied to the two spoken words, removing the phonetic detail whilst retaining the prosodic information, so children were parsing only the differences in lexical stress as close as possible. There were 28 trials, with target words divided into those exhibiting dominant stress (out of 15), and those exhibiting non-dominant stress (out of 13). The authors used a word reading measure to represent reading ability, and controlled for phonological awareness.

The word reading task was had a significant positive correlation with the dominant stress (r = .208, p = .004) and non-dominant stress (r = .343, $p < .001$) aspects of the Aliens Talking Underwater task, and following this a backwards elimination multiple regression was conducted to assess the contribution of four predictors (age (years), phonological awareness, dominant and non-dominant stress tasks) with the word reading outcome measure. The final model consisted of the Age (years), phonological awareness and non-dominant stress task as predictors with a significant contribution (dominant stress was removed because it was insignificant). Together, the measures predicted 76.3% of the variance in word reading accuracy. Why might only non-dominant stress have been a predictor of word reading ability, over and above age and phonological awareness? Arciuli posit that it may be because the processing of polysyllabic words with non-dominant stress is a more challenging skill to develop, with dominant stress an easier pattern of word reading to develop, as such the children's success on the non-dominant stress measure may have been a better indicator of their reading ability within this model. As well, the predictors of non-dominant stress and phonological awareness could have shared variance with the dominant stress task, incorporating this skill into their relationship. It is interesting to note that this is the only piece of research I am aware of that has used the Aliens Talking Underwater task, despite the potential displayed in the current study for this task across age groups. Compared to the tasks in Holliman et al. (2008) and Holliman et al. (2010a), the Aliens Talking Underwater task was designed to be appropriate for children in that it included a back story to keep the children involved and engaged, which is something much research does not account for in designing tasks with children, especially those who are pre-readers or early into their education. I will discuss some more studies which have tried to create child-friendly tasks for engagement in Section 2.7.

Lin et al. (2018) continue the trend of the previous studies, but used a different method to look into the prosody-reading relationship. They chose to adapt a task known as the Lexical Decision Task (henceforth known as the LDT), a measure of the implicit awareness of stress patterns. Originally designed for use with adults, Lin et al. adapted the measure for use with children. The stimuli were a mixture of di- and trisyllabic words, selected to ensure they were suitable for the children being worked with (aged between six and ten years of age). Nonwords were created by either changing the stress of a real word (e.g. CAbin becomes caBIN), which probed prosodic sensitivity, or changing the phoneme in the real word (e.g. cabin becomes calin), which probed phonological awareness. There were equal amounts of stress-shifted and phoneme-shifted nonwords, to test each skillset. All the children had to do was listen to the stimuli and decide whether what they heard was a real word or nonword. The researchers measured both the accuracy of the children's answers (real word accuracy, phoneme-changed accuracy, and stress-changed accuracy), and the response time to identify each word/nonword correctly. As well as this, a receptive vocabulary task and phoneme awareness task were included as control measures, and as the other tasks in this section have done, they focused upon word reading ability. Participants in this study spanned a range of ages; 22 six year old children, 22 eight year old children, 24 ten year old children, and 24 adults were also measured. The contribution to word reading ability was explored in a series of hierarchical regressions for each age each group, as a chance to investigate age-related changes in ability that may occur. Lin et al. found that only the six-year-old group of children showed stress sensitivity as a unique and significant predictor of word reading, after controlling for receptive vocabulary and phoneme sensitivity, as well as in a model accounting for receptive vocabulary and phoneme awareness. Unfortunately, the other age groups did not provide the same outcome, neither the eight-year-old or ten-year-old groups found stress sensitivity to account for significant variance in word reading ability, after accounting for each of the other control predictors. Interestingly, phoneme awareness was an important predictor to word reading throughout the age groups, suggesting that decoding skills continue to be influential to word reading throughout development. Stress sensitivity only being uniquely predictive in the youngest group of participants is interesting, and suggests the possibility that early decoding and vocabulary learning may rely more upon the ability to place the correct stress in words to learn and develop their internal repository of words. Over time this may become less important than phoneme awareness and vocabulary, as it becomes more about accessing the correct orthographic representation than decoding and placing the correct stress in the word. One weakness of the current study is the low number of participants in each group, with only 22-24 participants for each age group the conclusions are likely much weakened, especially when trying to compare groups with each other. However, as a potential new measure of prosody alongside phoneme sensitivity this project shows that looking at suprasegmental and segmental aspects of phonology together may be important to understand developmental differences in reading ability over time. The LDT is the first measure discussed in this section that directly contrasted phonemic and stress manipulations into a single study, and could be interesting to follow-up on.

2.5.2. Prosody and Multisyllabic Word Reading

So far, it seems that prosody is a potentially strong predictor of success in word reading ability. However, as this research has continued onwards the why of the relationship between prosody and word reading has been more closely discussed, with thought being put into whether prosody is more integral to the processing of multisyllabic words than monosyllabic words. I've discussed the importance of stress understanding and manipulation prior, but a challenge in reading multisyllabic words specifically is the application of lexical stress (Wade-Woolley and Heggie, 2015). Take for example the words *crocodile* and *giraffe*, the first has major stress placed on the first syllable (CROcodile), the second on the second syllable (giRAFFE). Readers need to bring this knowledge to the English orthography, as these features are not explicitly marked in writing. That is not to say there are not rules within English, aspects such as the change of morphology have consistent rules. For example, a neutral suffix like *-ness* does not shift the lexical stress from HAppy to HAppiness. Similarly, nonneutral suffixes such as *-ity* do lead to changes in stress (e.g. ACtive - acTIVity). But, there may be a contribution from prosodic understanding to multisyllabic word decoding, more so than monosyllabic words which are stressed the same way each time.

There are some studies which have put focus on this aspect of word reading in the literature. Wade-Woolley and Heggie (2015) conducted a study investigating if this understanding and awareness of stress difference in multisyllabic words could be predicted from prosodic and morphology ability with adult readers, and found this to be the case in that project. This was followed by Wade-Woolley (2016) the following year, using the same measures to investigate how prosody interfaces with multisyllabic word reading, with a focus on phonemic awareness as the control measure in this relationship, and working with children. The children were in grades four and five (mean age 10 years and 5 months) in Canada. The word reading task was created by selecting 33 words split into groups; 11 one-syllable words, 11 three-syllable words, and 11 four-syllable words. They were similar in frequency, and could also be divided into short (one-syllable) and long words (three- and four-syllable). Furthermore, a nonword reading task was also included, where children had to read 22 one-syllable nonwords from a list. The prosody task was named the Aural Stress Assignment task, a task designed to focus on participants' sensitivity to prosody at the word level. Children were asked to listen as words of varying syllabic length were played, and the children were asked to repeat the word twice, and then clap at the point where they heard the "main beat"'(i.e. the main stress in the word). This focused in on children's awareness of stress in multisyllabic words, and whether their pronunciation matched their perception of the stress placement. This was clearly designed to match the word-reading task in some way, the awareness of stress assignment will be intrinsically linked in some way to whether children can read the multisyllabic words presented.

In their results, Wade-Woolley divided the word reading stimuli into the aforementioned Short Words and Long Words grouping, to investigate the relationships of each grouping of words. Prosody was significantly correlated with both the short word ($r = .28$, $p < .001$) and long word ($r = .29$, $p < .001$) measures. For the regression model of long word reading, Wade-Woolley began with the predictors of prosody and phonemic awareness, and found that both measures provided significant variance to this outcome, supporting the idea that prosody would make a unique contribution to long word reading, with prosody providing a significant 4% additional variance to the measure. The authors remodelled this adding the nonword decoding task as an extra predictor, and found that phonemic awareness was no longer a significant predictor of long word reading, and prosody's unique variance was reduced to 2%, though still a significant predictor. The next model investigated prosody and phoneme awareness as predictors of short word reading. Both were significant predictors, phoneme awareness explaining 21% of the variance in total, with prosody explaining 4% of this. Though this project is interesting, it has a major weakness in only controlling for phoneme awareness (and nonword decoding in one model). Especially with the older age group that were worked with, there is so much that goes into reading - morphological awareness, vocabulary knowledge, working memory - that could have been included as control measures, and potentially would negate the relationships discussed here. However, this would have required many more participants to instigate, and is a common difficulty when trying to design research working with school children.

Wade-Woolley was also a part of the next study to be discussed, by Holliman et al. (2017b) who also investigated the relationship between prosody and multisyllabic word reading. This project is an improvement on the previous study, in that it includes more control measures as predictors of reading ability, namely; phonological awareness, receptive vocabulary knowledge, short-term memory, and morphological awareness. The authors worked with 50 year three children (mean age eight years) from a primary school in the UK. Prosody was measured with the DEEdee task, which was designed by Whalley and Hansen (2006). This is another stress sensitivity task, and involved children being presented with a phrase (e.g. Peter Pan) spoken by an English speaker. They then heard two more phrases, except the words were replaced by phrases made up of the word dee, stressed in varied ways (e.g. “DEE-dee-DEE" vs. “dee-DEE-dee"). This was presented under a forced choice paradigm, and children selected the DEEdee pattern they thought matched the original phrase they heard. This was of particular relevance for multisyllabic word reading, being a measure that investigated understanding of stress assignment in words, it was thought this may be a task that would have strong predictive ability for the word reading task. Furthermore, the task contains no phonemic content, being that children have to identify the stress patterns from DEEdee patterns, further separating the task from the phonological awareness measurement. The word reading task, named the Multisyllabic Word Reading Task, was used to assess children’s word reading accuracy, and also the frequency of different decoding errors made throughout the task. Throughout this section, the word reading measures have commonly been simple behavioural tasks, wherein children were asked to read a list of words which increased in complexity, as a representation of reading ability. For this project, Holliman et al. chose a task which goes into more detail about how well children could read these words. There was a list of 50 low-frequency multisyllabic words, ranging from two-to-six syllables in length, with low-frequency words used in particular to elicit errors in pronunciation, in the interest of seeing how the errors were related to the predictor variables. Errors were listed as so: decoding error, stress placement error, spondee error (pronunciation too slow to assign as one word), syllable error, or no attempt made. The frequency of these errors was totalled for each child. The task was operationalised as accuracy (the number of words recalled out of 50), and also as the individual errors types made.

To answer the question of whether prosody did make a direct contribution to multisyllabic word reading, the authors created a multiple regression model with all the aforementioned predictors as variables. Interestingly, it was both prosody (β = .454, t(44) = 4.298, p < .001) and morphological awareness (β = .290, t(44) = 2.496, p < .016) that ended up being significant independent predictors of multisyllabic word reading accuracy, and of particular interest was how strong the contribution of prosody was to this model. This is unlike the studies discussed prior, which may be due to the sole focus on multisyllabic words. Even more so than Wade-Woolley (2016), this study focused on using a multisyllabic word measure that was centred on making the participants use their prosodic skills to read the words. Part of the strength of this relationship could be the choice of the DEEdee measure, which focuses particularly on stress placement in words, which is an important aspect of multisyllabic word reading.

These two studies are the only projects I am aware of in the English language that focused specifically on the multisyllabic word reading aspect in relation to prosody. Both projects suggest that further research into this relationship is needed, to explore this over the time span of early reading development will be important also. For the study by Holliman et al. a focus was on children aged between seven and eight, but working with younger readers who have just entered into becoming literate the phonological awareness element may be more important as this aids early decoding, and older readers may benefit more from morphological knowledge which helps with more complicated words when suffixes and affixes shift stress.

2.6. Prosody and Reading Comprehension

A major aspect of the current project is the relationship between prosody and reading comprehension. As has been discussed, prosody is an important aspect of linguistic comprehension in speech, how a speaker uses prosody can affect the meaning the listener interprets. I've discussed the importance of the speech-reading relationship, and seen how word reading appears to benefit from prosodic ability, maybe even more so for multisyllabic words. However, as Groen et al. (2019) discussed in their paper talking about this relationship there has been relatively very little research into the contribution of early prosody skills to later reading comprehension. In particular, there are few studies investigating this relationship over an extended time period, and this area requires more work to explore this relationship. With this said, there have been studies that have looked at this relationship (Chan et al., 2020; Clin et al., 2009; Holliman et al., 2010b; Kim and Petscher, 2016; Lochrin et al., 2015; Whalley and Hansen, 2006; Whalley, 2017) which will be discussed throughout this subsection, and in the following subsection discussing multi component measures of prosody.

The first study to discuss is Holliman et al. (2010b), who investigated the relationship between prosody and reading comprehension in a study which looked to see if prosody was predictive of reading one year later with a group of children in years one, two and three in English schools. Prosody was measured via the aforementioned Mispronunciations Task in the previous section. The control measures included measures of phonological awareness, vocabulary knowledge, and the age of the participants. The authors conducted multiple models assessing word reading and reading fluency as outcomes measures in the regression models. Prosody (labelled speech rhythm in the paper) did not predict later reading comprehension ability, contributing a insignificant 2.3% of variance ($p = .057$). It may be that the use of solely the Mispronunciation Task is not the best representation of prosody and its relationship to reading comprehension. The task is a word-level measure, about identifying the correct stress pattern for these words. Comprehension will be affected by the incorrect stressing of words, as it may muddle a message being communicated. But, if related to the SVR model, the mispronunciation task may contribute more to the word reading/decoding aspect of reading development, rather than directly influencing reading comprehension. When looking at the word reading model prosody independently contributed 2.2% of additional significant variance, after controlling for the other variables. Though a small contribution, it demonstrates the potential skill Holliman et al. focused upon with the choice of prosody measure, and why the reading comprehension prediction was insignificant.

Another example of this relationship was conducted by Whalley and Hansen (2006), who investigated the relationship between prosodic sensitivity and reading comprehension. One aim of the project was to see if the measure of children's prosody ability could predict reading ability (namely word reading and reading comprehension) when controlled for by a measure of phonological awareness, which as I have discussed has a strong relationship with reading ability. Prosody was measured using the DEEdee task, detailed in the previous section. The DEEdee task showed a significant positive correlation with the comprehension task (.383, $p < .01$), and when entered into a regression model showed significant predictive ability for the comprehension measure when controlling for phonological and prosodic measures (.077 R^2 change at step 4, and a .093 R^2 change at step 3, both significant at $p < .01$). They also controlled for the previously discussed Compound Nouns task. However, unlike the DEEdee task, when entered into the reading comprehension regression model the contribution was insignificant (.02 R^2 change at step 3, and a .004 R^2 change at step 4, both insignificant contributions). The compound nouns task did show a unique contribution to the word identification measure, a measure of word reading (.042 R^2 change at step 3, and a .028 R^2 change at step 4, both significant at $p < .05$), so it did show some relationship to reading ability. In a final hierachical multiple regression model, Whalley and Hansen controlled for word identification, phonological awareness and non-speech rhythm at step 1, and then entered in the prosody measures at the following steps. The compound nouns task was added at step 2, and provided a very small amount of non-significant variance (R^2 = .001). The DEEdee task on the other hand did provide a unique contribution to reading comprehension after controlling for all of these variables (R^2 = .050, $p < .05$). The authors concluded that one measure of prosody contributing unique variance to reading comprehension provided good evidence that prosody may well be important to comprehension understanding, and that further work was required to explore this further. This early evidence that a behavioural measure of prosody could potentially be a unique contributor to reading comprehension was inspirational to research going forward, as will be seen through the rest of this subsection.

One of the authors of the Whalley and Hansen (2006) paper, Karen Whalley, also completed their PhD thesis exploring the role of prosodic skills in reading comprehension (Whalley, 2017). Prosody was represented by multiple tasks throughout the thesis, including the aforementioned Mispronunciation Task and DEEdee measures. They also implemented a task known as the Prosodic Nouns task, which required the children to listen to utterances and decipher where the prosodic mark was to differentiate the phrases, which could be three words (e.g. *sun flower pot*), a left branching compound noun/noun phrase (e.g. *sunflower pot*), or a right branching version (e.g. *sun flowerpot*). This is similar in design to the previous Compound Nouns task, but this task was a redesign to incorporate more prosodic modifications in this design. Last in the prosody battery was the Derived Word Production task (henceforth referred to as the DWPT). The DWPT is a morpho-phonological task, combining both prosodic and morphological elements into one design. Children had to attach a suffix to an English word (morphology) which in doing so may change the stress pattern of the word (prosody). For example, going from BEAUty to BEAUtiful does not alter the stress pattern, but ACtive to acTIVity does. These four measures of prosody provided a variety of prosodic abilities in the test battery, and gave Whalley ample chance to look into the details of what aspects of prosody may contribute to reading comprehension.

An important aspect of this thesis that influenced the current study was the inclusion of measures which represented the previously discussed Simple View of Reading. Decoding ability was represented by the use of two word reading tasks, one measuring single word reading accuracy, the second a measure of non-word reading accuracy. Linguistic comprehension was measured using a listening comprehension task, known as the Test for Reception of Grammar Version 2 (TROG-2 for short). The TROG-2 was designed by Bishop (2003), and presents children with short sentences which increase in grammatical complexity, and children have to listen and select a picture from a choice of four that they believe represents the sentence they heard. Whalley adapted this task, by selecting the sentences from the test that she felt would tap into prosodic ability for the participants. For example, the sentence *The cup but not the fork is red*, can be understood through the use of chunking words together (*The cup / but not the fork / is red*) and the emphasis of key words (*cup* and *fork*) to come to the conclusion of which picture is correct. The reading comprehension measure for the study was a modified version of the TROG-2, changing the placement of the nouns in the original sentences, as there were pictures which represented this altered sentences for each of their stimuli. This was important, as having the only difference between the linguistic and reading comprehension measure being the interaction with the task (listening to a sentence vs. reading the sentence aloud) is hopefully meaning the tasks are occurring in parallel, and focus in on the differences between these aspects of comprehension.

Whalley (2017) found similar results through the implementation of a number of studies with children, finding that prosodic sensitivity was an independent predictor of reading comprehension with her grade three children (Mispronunciation task: β = .22, R^2 change = .03, p > .05; Derived Word Production: β = .35, R^2 = .06, p < .05) when controlled by the components of the Simple View of Reading, but not for the older children in grade five (Derived Word Production: β = .18, R^2 = .02, p > .05), wherein Whalley suggests prosodic cue's direct impact on reading comprehension could have become less pronounced over time.

Whalley (2017) followed this up by investigating prosody and reading comprehension with grade four children, with the use of slightly different measures. Reading comprehension was represented by the Relative Clause task, similar to the TROG-2 task but using only the sentences that she believed required the children to use their prosodic skills to decode, in an attempt to maximise the contribution of prosodic sensitivity skill use. Again, this was matched with a listening comprehension version of the task. Prosody was again represented by the mispronunciation task (though slightly shortened), the derived word production task, and an extra task called the question/statement task, wherein sentence level prosody was assessed by deciding whether a spoken sentence was a statement or a question. Further to this, Whalley included a measure of morphological awareness separate from prosody, to see its independent contribution in the regression model. The final hierachical regression model with reading comprehension as the outcome was created, with word decoding, listening comprehension, and the measures of morphology entered for the first two steps as control measures. The derived word production and question/statement tasks were added alternatively at steps 3 and 4, but not the mispronunciation task which showed no significant correlation with reading comprehension. The overall model accounted for 30% of variance in reading comprehension, with each prosodic measure providing unique variance (derived word production task: Step 3 = 8% variance / Step 4

= 7.1% variance; question/statement task: Step 3 = 10% variability / Step 4 = 9.3% variance). The measures of morphology produced no significant variance in the model.

Prosody seems to have a predictive contribution to reading comprehension when controlling for the components of the SVR. This was a direct influence on the current study, as apart from Whalley's thesis there is no other project I am aware of that investigated the prosody and reading comprehension relationship whilst explicitly controlling for the aspects of the SVR. This was a gap in the research that I took on to help further explore too. As well, the possibility that prosody ability could independently contribute to reading comprehension over and above these components is a point to consider. Should prosody be considered another aspect of the SVR framework? And, what else could be missing? I reflect on these questions in Chapter 5.3. For now, I see the potential for prosody being an independent predictor of reading comprehension, but these were not the only studies.

Another project which investigated prosody and reading in presence of word reading and linguistic comprehension was conducted by Kim and Petscher (2016). The researchers worked with 370 first-grade children creating multiple models to explore the relationship between prosody and word reading, and prosody and reading comprehension. Focusing particularly on prosody and reading comprehension, the authors explored whether prosody's relationship to reading comprehension was completely mediated by the presence of the word reading and linguistic comprehension measures. I will discuss mediation analysis in more detail in Chapter 2, but mediation is a modelling technique by which researchers can use many linear regressions to assess whether an independent variables relationship with a dependent variable is dependent upon a third hypothetical variable, namely the mediator. In the context of the study, Kim and Petscher wanted to explore whether the relationship between prosody and reading comprehension was actually dependent upon the children's knowledge and capability of word reading and linguistic comprehension (i.e. the variables of the SVR). To do this they created multiple confirmatory factor analysis models to explore prosody's relationship in the presence of these variables. Prosody was measured using the word stress task designed by Holliman et al. (2012), wherein children heard a word and were asked to identify the stressed syllable they heard by pointing to a box representing one of the syllables. Reading comprehension was measured by a passage comprehension task where the children filled in the blanks in short passages of text, and a more traditional passage reading task with comprehension questions. The authors compared two possible models with the reading comprehension outcome, this was essentially whether prosody was completely mediated by word reading and linguistic comprehension (Model 1), or a direct path between prosody and reading comprehension was allowed after accounting for the relationship through word reading and linguistic comprehension (Model 2), a partial mediation model. The outcome of this modelling was that Model 1 was a better for the data. 92% of the variance in the reading comprehension construct was explained by the predictors, so this model was a particularly excellent fit. Focusing on the regression lines they found their measure of prosody was weakly, but significantly related to both linguistic comprehension (.20) and word reading (.29), and showed significant indirect effects with reading comprehension (.28, $p < .001$). This model was important to the current project, as it suggested that there was an interface between prosody, reading comprehension and the SVR components. One critique of the study is that the measure of prosody was solely for the stress component of prosody, and as previously discussed the definition of prosody expands beyond simple stress perception. This influenced the choice of prosody measure for the current study, which I discuss towards the end of this section.

Prosody and Reading Comprehension with Morphological Awareness as a Predictor

The studies above explored the prosody-reading comprehension relationship, often with a measure of phonological awareness as a control in the modelling and data analysis. However, there must be other aspects of language that are related to reading ability. One of these aspects is morphological awareness, that is awareness of the internal structure of words, and the ability to understand and manipulate this structure (Clin et al., 2009). For example, knowing that you can take a stem such as *friend* and affix extra parts to add extra meaning, such as *friend + ly*. As the current project is not about morphological awareness, I do not spend much space going into detail about this aspect of language ability, but there is research in the literature that has included morphology in their exploration of prosody and reading as a control measure, and it will be seen if it affects this relationship.

The first of these is Clin et al. (2009), who worked with 104 8-13 year old's to explore how prosodic sensitivity, morphological awareness and reading ability related to one another. The study used two tasks to measure prosodic sensitivity, the first being the Stress Contour Discrimination task adapted from Wood and Terrell (1998). This task involved children listening to two sentences, firstly without distortions, then with a low-pass filter applied to the sentence removing the phonemic information from the speech, whilst retaining the prosodic contour. This is alike to the Aliens Talking Underwater task discussed above, in which children had to use their knowledge of prosodic contours to tell the researcher whether the low-passed sentence was the same as the normal sentence. The second task was the previously discussed DEEdee task, a measure of stress contour matching. Reading ability was represented by word reading measures from the Test of Word Reading Efficiency (TOWRE) designed by Torgesen et al. (2012), to score real and nonword reading ability, and the Word Identification subtest from the Woodcock Reading Mastery Tests (Wood and Terrell, 1998). Reading comprehension was measured by the Gray Oral Reading Test (Wiederholt and Bryant, 2001), which involves reading a passage of text which is measured for aspects of fluency and rate of reading, followed up by comprehension questions on the text that was read. Lastly, morphological awareness was measured by a task which involved children producing the derived form of a given stem (e.g. appear-appearance, major-majority). In analysing the relationships, the authors created a composite of reading ability, which combined the word reading and passage reading tasks into one variable. This is important to acknowledge, as despite being one of the few studies to investigate reading comprehension, Clin et al. integrated this measure into a single composite measurement. The issue is that when the models are completed, the researchers cannot then extrapolate which predictors were particularly strongly related to word reading and which to comprehension, as despite both being reading ability it is likely they require differing skill sets to achieve. This is a limitation of the study.

In their correlations they found that their measures of prosodic sensitivity (also a composite variable) to be significantly correlated with reading ($r = .56$, $p < .001$), and reading was correlated with morphological awareness (stress-neutral morphological awareness: $r = .70$, $p < .001$; stress-shifting morphological awareness: $r = .76$, $p < .001$). Prosodic sensitivity was also significantly positively correlated with the measures of morphological awareness (stress-neutral morphological awareness: $r = .53$, $p < .001$; stress-shifting morphological awareness: $r = .48$, $p < .001$). In their linear regression model, they entered the predictor variables step-by-step. The first step was their control variables (age in months, nonverbal intelligence, general language ability, and phonological awareness). Alone these variables contributed 59% of variance in the prediction of reading ability outcome. When adding prosodic sensitivity and morphological awareness they found slight R^2 changes, with morphological awareness providing a larger significant contribution than prosodic sensitivity (Prosodic sensitivity: Step 2 - R^2 change = .03 / Step 3 - R^2 change = .02; morphological awareness: Step 2 - R^2 change = .07 / Step 3 - R^2 change = .06). The models showed that prosodic sensitivity provided an independent contribution to the variance of reading ability, but this may have been weakened in relation to morphological awareness. However, the creation of a composite measure of reading ability prevents us from learning about the differentiation between word reading and reading comprehension relationships, which would be interesting to know in regards to these variables. The next study I discuss allows us to explore this further, with reading comprehension as an independent measure.

If this composite measure was an issue in disentangling the prosody-reading comprehension relationship, then the next studies to discuss rectify this issue. First is Deacon et al. (2018), who investigated the unique contribution of morphological awareness and prosodic sensitivity with measures of reading ability, in this case word reading and reading comprehension. An important note to make about this project, is that unlike a lot of the studies discussed there is a longitudinal element to this project. The data collection was first conducted with children when they were five and six years of age (school years one and two), and is detailed in a research paper by Holliman et al. (2014a), where they were collecting data to test a new measure of prosodic sensitivity, the Dina the Diver study (discussed in more detail in the next subsection). They collected data about children's prosodic sensitivity skills (the Dina the Diver task), phonological awareness (rhyme and phoneme awareness), vocabulary, and morphological awareness. Holliman et al. worked with 75 children at this time, and in the follow-up study by Deacon et al. 70 of this original group took part. At this time the group of children had a mean age of eight years and six months (school years three and four), and they completed standardised assessments of word reading, passage reading accuracy and reading comprehension as the measures of reading ability. With this lapse of time it is interesting to see the long term contribution of these metalinguistic skills to later reading ability, and if they are predictive of reading.

To investigate these relationships the researchers conducted a multiple regression analysis, with each of the reading ability variables as predictors. I am particularly interested in the reading comprehension model, which included the variables phonological awareness, vocabulary, prosodic sensitivity, morphological awareness. Furthermore, there was a version of the regression model where word reading was also included a control measure. Unlike the above study by Clin et al. (2009), Deacon et al. did not find a unique contribution of prosody to reading comprehension when entered into the model without the word reading control (β = .045, R^2 change = .001, p > .10) or with the word reading control measure (β = .004, R^2 change = .001, p > .10). Morphological awareness on the other hand, provided a significant change in variance without (β = .467, R^2 change = .108, p < .05) and with word reading as a control measure (β = .383, R^2 change = .066, p < .05). It appeared from their study that prosody may have been subsumed by morphology in its independent contribution to reading comprehension over a longitudinal period. In discussing why this relationships exists as it does, the authors reflect that prosody's contribution may come through in the reading of multisyllabic words, and may be more important earlier in reading development when these words are harder to read. They suggest that prosody's contribution to reading ability may ebb over time, becoming more of a supporting skills to other aspects, in this case morphological awareness.

The final study in this subsection, conducted by Chan et al. (2020), continues the theme of investigating prosody as a predictor of word reading and reading comprehension ability, whilst also including morphology as a control measure. The authors highlighted that the relationship between prosody and morphology comes with the reading of multisyllabic words with suffixes such as "-ity" or "-tion", being that they change where the primary stress appears in the word between the original and derived word with the suffix. The study worked with children in Grades Four and Five (ages 9-11) to explore this relationship. They hypothesised a low correlatory relationship between the measure of prosody and reading comprehension based on the previously discussed Kim and Petscher (2016) study. Prosody was measured using the same method as in Wade-Woolley (2016), identifying the main syllabic stress in multisyllabic words. Reading comprehension were measured via two methods, similar to Kim and Petscher (2016) they were a passage comprehension task to fill in missing words in a text, and a more standard passage reading task with comprehension questions following. Focusing on the reading comprehension outcome measures, the authors found that their prosody measure was not a significant contributor to reading comprehension in their hierarchical regression analysis, when entered before or after morphological analysis. The authors suggested that prosody may have shared variance with other predictors that support reading comprehension, as it has been significantly correlated prior to this modelling. This was supported by their mediation model, where when accounting for morphological awareness, word reading and phonological awareness prosody showed a non-significant relationship with reading comprehension, only indirectly related via morphological awareness and word reading for an indirect effect of .22. The important takeaway from their work was that prosody was indirectly related to reading comprehension through word level processes (word reading and morphological awareness), which in turn support the development of reading comprehension. Could it be that prosody plays a supporting role to the word reading, i.e. decoding? This outcome may have been come through the use of a prosody measure focused on stress sensitivity with word-level stimuli. Depending upon the measure and stimuli used this model could have had a different output, and again this inspired the current design of the project and the choice of prosody task used.

Prosody and Reading in Non-English Languages

The focus of the current thesis is on prosody in the English language, but it is important to reflect on work that has been conducted exploring prosody in other languages and what we can learn from this. Importantly, whether other languages see a similar result for the contribution of prosody to reading comprehension. Below is a short description of some key studies that were considered thinking about this issue.

Many of the studies read were from work with Spanish readers. In comparison to English, Spanish is an interesting language to consider for prosody, as the written language includes stress marks to indicate the appropriate stress to make on the word. The expectation would be that those learning to read Spanish would be aided by this inclusion, and as such the relationship between prosodic understanding and reading comprehension would also be strengthened. The first example comes from González-Trujillo et al. (2014), who work with 66 third grade children (ages 8-9) to investigate whether prosody (measured as nonspeech rhythm skills) aided reading acquisition, expressed by word reading in this study. They controlled for verbal intelligence and working memory also. The results found a high correlation between the measure of decoding and word reading, which is to be expected. However, nonspeech rhythm was also highly correlated with all measures of word reading ability (reading fluency, $r = .42$, $p < .01$; stress reading, $r = .34$, $p < .01$), and in a hierachical regression analysis also predicted unique variance in reading fluency ($\beta = .226$, R^2 change = .042, $p < .05$) and stress reading ($\beta = .282$, R^2 change = .065, $p < .05$). The authors concluded that the results showed nonspeech rhythm as having a strong relationship with Spanish reading, and this may be a universal factor in reading acquisition, as they earlier reflect on the differences between the Spanish and English language as a discussion point. A similar project by Defior et al. (2012) looking at stressed and unstressed syllable awareness found that stress awareness also has a strong independent relationship with word reading ($\beta = .433$, R^2 change = .165, $p < .05$). This same outcome has been seen in many Spanish studies (Gutiérrez-Palma et al., 2016, 2009; Gutiérrez-Palma and Reyes, 2007), suggesting that the relationships seen in the English studies is not unique.

Veenendaal et al. (2016) was a study carried out in Holland following 99 Dutch primary school students investigating the relationship between the segmental and suprasegmental aspects of phonology and their contribution to reading comprehension. It was a longitudinal study that followed students through grades four and five measuring their segmental and suprasegmental skills, and then added reading comprehension as a measure in grade six. They used structural path modelling to investigate the relationships, and they found that suprasegmental prosody had some predictive power for reading comprehension one year later, above that of the decoding measures in the study. It is important to remember however that Dutch is a language with high orthographic-phonological consistency, unlike English, which may explain the ability for the suprasegmental measures to mediate between decoding and reading comprehension at year six. In English this same result may not have been as strongly defined, and why decoding measures remain strongly related to reading comprehension late into children's reading development. What these examples demonstrate is that despite English being a more opaque language than others described here, there is a similar pattern of prosody contributing to reading ability.

Summary

The previous subsections of this chapter have been attempts to understand the relationship between prosody and reading ability. I began discussing prosody's relationship in studies using word reading as the reading measure. The majority of the studies used a measure of prosody which required children to identify or understand stress in some way, likely chosen due to the importance of stress sensitivity in reading words. In general, prosody seemed to contribute to word reading ability with the younger child participants, but less so with the older participants. It may be that the older children have internalised their ability to read words, through experience of reading more difficult texts, as well as other linguistic skills maybe taking over, such as morphological awareness. Furthermore, I had a short discussion regarding studies which had focused specifically on multisyllabic words, which varied in the strength of the contribution, but found that prosody was a significant predictor of multisyllabic word reading ability. This may be an avenue for future research, to determine if prosody is an important aspect, or mediated by skills such as morphological awareness.

This was followed by looking at reading comprehension, the other aspect of reading ability. The outcome of this relationship is more inconclusive, with some studies displaying some contribution to reading comprehension ability (Clin et al., 2009; Kim and Petscher, 2016), whilst others did not (Chan et al., 2020; Deacon et al., 2018; Holliman et al., 2010b), or were mixed depending on the measure (Whalley and Hansen, 2006) or age of the children worked with (Whalley, 2017). This variation in outcomes was interesting, with many authors concluding that it may be that prosody mediates or is indirectly contributing to reading comprehension ability through other linguistic skills, such as morphological awareness of word reading. This seems plausible, as morphological changes of words require stress awareness to pronounce the new word correctly. As well, word reading, as seen, is somewhat influenced by children's prosodic understanding. It just may be that prosody isn't as important as these other skills. There is an issue though, each of these studies either relied on a behavioural measure of prosody that was about stress awareness, even though prosody is wider in definition than this. This leads into the next subsection, discussing studies which have either designed or used a multi-component measure of prosody, to investigate many representations of prosody at once.

2.7. Multi-Component Measures of Prosody and the Reading Relationship

An important aspect of prosody is the complexity of prosody, and how it can be defined by so many aspects, whether that be stress, timing, or intonation. To understand the prosody-literacy relationship it would be beneficial to use a measure that incorporates multiple aspects of prosody (such as the aforementioned stress, intonation and timing definition), as well as investigating these aspects at different linguistic levels (such as word level, phrases, sentences). Prosody does not have such tests, or at least tests that are used with regularity to be comparable. Through this discussion, prosody has been represented by single measures, but what measures in the literature are there that attempt to measure multiple aspects of prosody? In this section I discuss studies which have used multi-component measures of prosody to explore word reading and reading comprehension.

The first multi-component measure to discuss is the Dina the Diver multi-component task, designed and tested in two papers by Holliman et al. (2014a,b). The Dina the Diver measure was an attempt to create a measure of prosodic sensitivity that allowed for the measurement of multiple aspects of prosody at different linguistic levels. The measure was developed to be used with children, and had novelty in that there is a character, named Dina, who guides the children through the task.

- The first task represented stress sensitivity, and involved the character producing utterances ‘above water’ in a clear voice (e.g. Winnie the Pooh), and then hearing two utterances low-pass filtered with similar constructions, but differing stress patterns (WInnie the POOH vs. HUMpty DUMpty). Children would listen to these low-pass utterances and then through a two-forced choice paradigm select which utterance was the same as what they originally heard.

- The second task assessed sensitivity to intonation changes, where child heard an utterance and has to say whether the character was telling them something or asking them a question (e.g. ‘Aladdin’ vs. ‘Aladdin?’).

- Lastly, there was an assessment of sensitivity to syllabic timing, in which children heard an utterance spoken "under water" twice, and they had to identify whether the two utterances differed in initial syllable duration (e.g. Tinkerbell vs. Tiiiiiiiinkerbell is an example of utterances that did not match).

With these three tasks, the hope was that the measure could cover multiple components of prosodic sensitivity, and that this task would allow for the measurement of a wider variety of children’s prosodic skills. In testing this measure they looked at the correlations between prosody (as a composite measure) with the other factors they had measured. Within the research paper where they explored this measure (Holliman et al., 2014b) they saw that the measure of prosody was correlated with their measure of reading comprehension ($r = .42$. $p < .01$), a passage reading task with comprehension questions. Unfortunately, the focus was not on the assessment of the prosody-reading comprehension relationship in this instance, as the project was an exploratory study to assess the Dina the Diver task. The partner research paper (Holliman et al., 2014a) is also not a direct assessment of reading comprehension ability, rather a test of a model to explore prosody's relationship with speech and reading skills (see Subsection 2.3 where I discuss this paper). Inspired by this measure, there is also have the aforementioned 'Aliens Talking Underwater' task by Arciuli (2017), which focused purely on stress perception, but using the low-pass filter technique. Unfortunately, it was not a multi-component measure, and as such this measure has only been implemented into a few studies of prosody and reading ability (Deacon et al., 2018; Holliman et al., 2014a,b).

There was an attempt to re-design this measure, first described in Holliman (2016), called the Brenda's Animal Park task. Children were asked to help a character named Brenda in their animal park with four different tasks, each designed to represent a feature of prosody. Each of these tasks was pooled into a composite measure of prosodic sensitivity.

- The first task was a measure of sensitivity to compound nouns, that is the use of speech rhythm (e.g. children were asked to decide if they heard a compound noun like "ladybird", or a noun phrase like "lady" and "bird").
- The second task was a word stress task, where children listened to a pre-recorded utterance of a word, with either the correct stress pattern (e.g. CROcodile) or with the incorrect stress pattern (e.g. croCOdile). Children were then asked to decide which of the if the utterance sounded right or wrong based on what they had heard.
- Third, was the intonation task, where children decided where a pre-recorded utterance sound like a statement or a question, similar to the intonation task described in Holliman et al. (2014a).
- Lastly, is the phrase stress task, wherein child had to decide which of two pre-recorded utterances (e.g. 'apple pie' [strong-weak-strong], and 'tomatoes' [weak-strong-weak]) matched a "Ba-BA" pattern e.g. if the stress pattern was BA ba BA (strong-weak-strong), then "apple pie" would match, but "tomatoes" would not. This task is similar to the DEEdee task by Whalley and Hansen (2006) discussed previously.

There are two examples of the use of this measure. The first, is the aforementioned Holliman (2016), who worked with children at two time points; time one they worked with 101 four and five year old children, and at time two 93 of these children remained and were between five and six years of age. They controlled for vocabulary, phonological awareness (rhyme and phoneme), morphological awareness, and measured these against word reading and spelling measures collected at time two. The correlation analysis showed this new measure of prosody (labelled speech rhythm sensitivity in this paper) was positively correlated with the measure of word reading ($r = .259$, $p < .05$). When entered into the multiple regression model, prosody became an insignificant predictor of later word reading ability, with phonological awareness the only significant predictor of word reading, explaining 10.5% of the variance in the model (prosody for comparison was measured to contribute .04% variance). Lastly, a mediation analysis was conducted to analyse whether, in predicting word reading, prosody's direct relationship was mediated by the other control variables. When entered alone, prosody was a significant predictor of word reading ($\beta = 0.259$, $t(89) = 2.531$, $p = .013$). However, when entered simultaneously with either of the control variables, this direct relationship was made insignificant, suggesting prosody was being mediated. Looking at the mediation model, prosody was mediated by its relationship with phonological awareness ($z = 2.82$, $p = 0.005$), which is in line with the multiple regression model. The results fit with previous research regarding the relationship between speech skills and early reading ability, that is children who are more sensitive to phonological awareness are more likely to succeed at reading tasks. This is likely why the measures of vocabulary, morphology and prosody were not as strongly linked to Year 1 reading ability. In terms of the new task, Holliman suggested that the prosody task was a success in its use here, and represented a potential new measure for prosody use with young children.

This task was also used in another piece of research by Holliman et al. (2017a). The Brenda's Animal Park task was used to represent a measure of prosodic sensitivity, and was joined by control measures of vocabulary, phonological awareness, and morphological awareness. To represent reading a measure of word reading was used, investigating the contribution of these measures to investigate mono- and multisyllabic word relationships, assessing where there was a difference in contribution from these skills. The researchers worked with children in Year 1 (ages 5-6) using a cross-sectional design. Prosody was found to have a significant positive correlation with word reading ability ($r = .39$, $p < .001$). For their initial multiple regression analysis they entered each variable in a stepwise procedure, with prosody entered at Step 4. The measure of prosody was significant predictor of word reading ability when entered at step, explaining 3.8% unique variance in word reading. The final stage of this study was the exploratory measure of monosyllabic and multisyllabic words as outcome measures. This is due to the idea discussed prior that prosody is likely to aid multisyllabic word reading, through processes like appropriate stress assignment in words. Prosody was once again added at Step 4 of the model, after the other control variables. Prosody was able to explain a significant amount of unique variance in both monosyllabic (3.8%) and multisyllabic (13.5%) word reading. Most interesting is the greater variance explained in the multisyllabic reading by prosody, fitting with the aforementioned thought that prosody was particularly useful for decoding words with multiple syllables, and adding to the previously discussed research. As above, the Brenda's Animal Park task appeared to be a successful representation of prosodic skill with young children, but has not been used beyond this project.

In the above list of tasks there are a number of authors attempting to create a new multi-component measure of prosody. This reflects the complicated concept of prosody, there is so many aspects you can chose to represent prosody, and these tasks fit with the previously described idea of their being three major aspects of prosodic understanding; stress, timing and intonation. These are the only projects I am aware of that have used these multi-component measures, or more specifically have designed a single task which incorporates aspects of prosody into once task. This leaves a gap in the literature to explore, which this study will attempt to fill. However, there is one more multi-component task to discuss which has been around since 2003, which has a direct influence on the design of the current project.

Lochrin et al. (2015) is the final study discussed in this section. The authors investigated the relationship between prosody and reading using a multi-component measure of prosody, but they chose to do this in a unique manner. In their study, Lochrin et al. split the concept of prosody into receptive and expressive prosody. Receptive prosody skill is the understanding or awareness of prosody use, such as understanding when a statement made by a speaker is a question or a statement. Expressive prosody on the other hand is the production of prosody, that is how prosody is used by a speaker to convey meaning. Lochrin et al. claim this study was the first exploration of receptive and expressive prosody and how they relates to three aspects of reading ability; accuracy of reading aloud words and nonwords, and reading comprehension.

The exploration of these two aspects of prosody was inspired by the method they chose to use for this research. The Profiling Elements of Prosody in Speech-Communication 2015 battery test, which is referred to going forward as the PEPS-C, was created by Peppé and McCann (2003), and was devised as a semi-automated test of prosody skills. Its primary use was for understanding the speech of children diagnosed with high-functioning autism/Asperger's syndrome. However, the test has been used beyond this to measure prosody skill in other contexts, as the battery of tests measures various prosodic forms and functions, a rarity in the measurement of prosody via behavioural tasks. The PEPS-C is not yet standardised, but its potential utility for research comes from the variety in the 14 tasks. Furthermore, the tasks can be split into two response types, non-verbal and verbal. These can be differentiated into receptive prosody and expressive prosody. Part of Lochrin et al. purpose in the study was to investigate the utility of the PEPS-C, and whether it could be a useful tool for examining the link between prosody and reading ability. This study was informed by previous research, much of which has been discussed throughout this chapter. The majority of previous research had been conducted investigating the link between receptive prosody and the reading of single words. They cited Whalley and Hansen (2006), Holliman et al. (2010a), Holliman et al. (2010b), and Goodman et al. (2010) as examples using receptive prosody for this purpose. Though there were was a prosody and word reading relationship, when phonological awareness was taken into account there was less variation seen. Comparatively, there is less investigation into the relationship between receptive prosody and nonword reading. They summarise Whalley and Hansen (2006) who found no relationship between receptive prosody and nonword reading, as well as Holliman et al. (2014b) who found a positive significant correlation between receptive prosody and nonword reading accuracy with younger children. These contradictory results led the authors to also look at this relationship in their study. These outcomes created the expectation that the receptive prosody tasks in the PEPS-C would be closely related to word and nonword reading accuracy. They also summarised the relationship between receptive prosody and reading comprehension from previously discussed studies (Holliman et al., 2014b, 2010a; Whalley and Hansen, 2006). As identified in Section 2.6 the findings were mixed, with two of the studies finding a relationship, and one no association.

Turning to expressive prosody, Lochrin et al. highlighted that little is known about the link between expressive prosody and word/nonword reading, and that this is an avenue very much unexplored in the literature. They cited more research regarding the relationship between expressive prosody and reading comprehension, focussing on research that has investigated oral reading fluency (e.g. Miller and Schwanenflugel, 2008). From this work, the authors suggested that children's ability to read with fluency (which includes prosody) indicates better reading comprehension. With that said, there is a difficulty in interpreting expressive prosody elicited by reading, as the enhanced expressive prosody could be a result of successful reading comprehension rather than a precursor to it. This was a major reason for Lochrin et al. using the PEP-C battery test, as this potential conflation between expressive prosody and reading ability could be somewhat reduced by the oral language tasks in the PEP-C. Rather than having children read words or sentences, the expressive prosody tasks elicited children to speak in response to a picture or another voice e.g. imitating a phrase with the same intonation. The lack of research made it difficult to be conclusive, but the authors predicted that expressive prosody was more likely than receptive prosody to be related to reading comprehension, due to children using have to use appropriate prosody in their speech, showing understanding of prosody use rather than just understanding.

The authors worked with sixty-three children in Sydney, who were aged between seven and twelve years of age, with English as their first language. The final results fit somewhat with their initial hypotheses. Measures of receptive prosody were indeed significantly related to word and nonword reading, when placed in a multiple regression model of each, with some measures of expressive prosody also significant predictors. Of more relevance to the current project was the modelling of reading comprehension with these predictors. This model provided three significant predictors; age of the children, a measure of phonological awareness, and one test from the expressive prosody set of tests (Expressive Chunking). Together these predictors accounted for 40% of the variance in children's reading comprehension, with Expressive Chunking providing 15.44% of this variance total. It seems that expressive prosody, and by extension receptive prosody, may be a predictor of reading comprehension ability. The authors noted that they believe these predictors contribute to reading comprehension, indirectly through the measures of the aforementioned Simple View of Reading. It was this potential avenue for further research which prompted the current study to explore the relationship between prosody and reading comprehension, using the SVR framework as a guide to the measures to use. Lochrin et al. provide an example of this, suggesting receptive prosody may contribute to reading comprehension through its enhancement of decoding ability, such as the understanding of the use of stress in speech.

The study did show prosody contributing to reading comprehension, and suggests that exploring the PEPS-C as a representative measure of prosody ability is a viable option. I discuss the PEPS-C itself more in Section 3.2.1, but for now it is important to know that the idea of receptive and expressive prosody was important to the current study. It was also a gap I wanted to explore further, as the previous studies in this section (and across the chapter) mainly focus on receptive prosody tasks, and I wanted to distinguish between the contributions of receptive and expressive prosody.

It is important to note that Lochrin et al. indicated some issues with the PEPS-C when completing their research. However, despite these issues, the authors say that the PEPS-C is the most comprehensive test of receptive and expressive prosody in oral language currently available. Firstly, the set of tests do not cover every possible type of prosodic processing, though this is partly due to the complexity of trying to describe and measure such aspects of language. The authors provide the example of irony, a complicated aspect of meaning that is difficult to describe, and very difficult to try and consider how to operationalise and measure. Secondly, because PEP-C is a test of oral language ability, Lochrin et al. (2015) suggest that the PEPS-C could include spectrographic analysis of responses for the expressive tasks. Spectrographic analysis is a method of speech analysis that allows one to study the acoustic-phonetic characteristics of speech. This might be comparing pronunciations between speakers or looking closer at the spectral property of vowels or consonants used in speech. For investigating prosody, one can view a graphical representation of one's recorded speech, for example where a pause has been made in speech and could identify the length of the pause and assess the significance of this. The PEPS-C does include the ability to record children's speech as they complete the tasks, but has no internally designed feature to use this for any further data exploration. Finally, the authors suggest longer utterances could be included in the phrase-level tasks, as the current stimuli are currently 6-7 syllables in length. This design is likely due to the original audience for the PEPS-C being young children with autism, so the tasks have been designed to be appropriate for this audience.

Though it can be improved it stands as a unique and useful tool for being able to explore these aspects of prosody. This was why the current project uses the PEPS-C as the measure of prosody, it provides a comprehensive battery of tasks for analysing prosody, It also helps to explore the gap in the literature exploring the differential contribution of receptive and expressive prosody to reading ability.

Summary

Because of the variation in ages, methods and modelling techniques there is no conclusive answer of prosody's contribution to reading comprehension. Whalley and Hansen (2006), Lochrin et al. (2015) and Whalley (2017) each suggested to varying degrees that prosody contributed to reading comprehension, even in the presence of predictors shown to have previously established impact on reading comprehension (e.g. phonological awareness). Even a study such as Chan et al. (2020), which found prosody's independent contribution non-significant, but suggested that prosody could play a supporting role in this relationship to other aspects related more directly to reading ability. The prosody-reading comprehension relationship remains an area of research which requires more work. Part of this may be due to the focus of prosody measures on word-level prosody rather than over longer chunks of text, which is just as important, maybe even more important as speech being a continuous stream of information requires continuous prosodic changes. Important to note at this point is that prosody is multifaceted and hard pin a description onto in research, and one important avenue of research uses spectographic analyses to assess prosody use in speech, which could help to explore the relationship between expressive prosody and reading ability. I discuss this further in Chapter 5 in the limitations and future directions section, but suffice to say the current study is focused upon prosody as an aspect of speech measured through quantitative behavioural means, extrapolating meaning from the relationships within. The physical presence of prosody is important also, but was not explored in the current project.

2.8. Prosodic Skills to Aid Reading

In the previous subsection I reviewed and discussed the relationship between prosody and various aspects of reading ability. However, there are more prosody skills that are involved and investigated as regards reading ability. I provide a summary of two of these areas of research, prosodic reading and implicit prosody, to give an idea of the alternative ways prosody can be defined in a research design, as well as how to measure prosody.

2.8.1. Prosodic Reading

I have already stressed the importance of prosody in speech comprehension, which in turn aids reading ability in a reciprocal relationship. One aspect of prosody not discussed yet is its utility for oral reading. The role prosody plays in oral reading ability may be important for self-teaching and overall development of reading comprehension. Before discussing this relationship, I will describe the skill of reading fluency.

Reading fluency is the ability to read written text with appropriate speed, accuracy and a natural tone of voice. The ability to read text fluently is important for developing readers, as one of the methods for measuring children's reading progress in school is via oral reading. Adults spend the majority of their time reading silently, but if teachers are to assess children's progress they need to hear the children read aloud. Furthermore, beginning readers may come to understand and develop their reading ability by being able to listen to themselves read (Kuhn et al., 2010). That is, children comprehend the text through their own oral reading ability, and the more fluent this is, the better able they will be able to take meaning from the text they read. Early oral reading begins in quite a staccato voice, children haven't developed the ability to decode the words on the page, understand the meaning of these words together, and use the appropriate voice to say these words out loud in tandem. As with all aspects of reading, reading fluently is complicated and takes effort and time to achieve. It also requires teaching children that good reading is not just reaching the end of the text as quickly as possible (Godde et al., 2019), rather it is to be fluent and clear to ensure that the listener successfully understands what is being read.

An important aspect of successful reading fluency may be children's prosody use. If prosody is the melody of language, then it stands to reason that proficient use of prosodic ability would create a fluent reader. Calet et al. (2019) list prosody as an aspect of fluency, alongside the aforementioned speed and accuracy of reading. There is also the idea of “prosodic reading" in the literature, a prosodic reader being one who utilises aspects of prosody (appropriate pauses, paying attention to syntax and punctuation, using intonation to improve comprehension) to better comprehend a piece of text. It is a definition that lacks precision according to Godde et al. (2019), as it may be that prosodic reading is just being a good oral reader? Groen et al. (2019) debated this in their paper, as text reading prosody and reading comprehension both depend on the ability to read well as a base, so is it possible that prosodic reading is an epiphenomenon of just being better at some of the base skills of reading ability? For example, I have discussed decoding ability as a precursor to reading comprehension, and a foundational skill for children to become literate. It may it be that children who are able to read words with automaticity are poised to read better, and that prosodic reading is children becoming more proficient with reading in a general sense, and not a separable skill for measurement. This is an issue throughout reading development, attempting to separate a specific skill from the general improvement of reading ability. It may be that identifying aspects of what defines prosody - e.g. appropriate intonation, correct stress - as identified in 2.1 with the three components, this allows for the ability to summarise what in particular makes a good prosodic reader, focussing on these prosodic elements specifically rather than as a holistic concept.

There is an argument for the presence of prosody as an active and important part of reading aloud, particularly demonstrated in research by Paula Schwanenflugel and other researchers she has worked with (Benjamin and Schwanenflugel, 2010; Kuhn et al., 2010; Miller and Schwanenflugel, 2006, 2008; Schild et al., 2014; Schrauben, 2010; Schwanenflugel and Benjamin, 2017). Each of these studies was in some way an attempt to explore how expressive prosody skills are related to successful fluent reading, and how this in turn could aid reading comprehension. Important to note, is the way prosody was measured in these projects, which is through the recording of the children's speech and analysing prosodic aspects of the stream. Taking Miller and Schwanenflugel (2008) as an example, prosody was represented by analysing recordings of an oral reading fluency assessment. The authors selected prosodic features they felt would be useful to distinguish readers reading fluency, such as pause duration, changes in the fundamental frequency (a representation of intonation change), and sentence-final pitch change. In turn, the authors were able to model the changes in these prosodic features between

Grades one and two, and then to examine word reading skill and oral reading fluency. By looking at the physical aspect of reading ability, the authors could mark out prosody explicitly and examine how the children were using it. Throughout this research correlations between prosodic reading ability and reading comprehension has been found with early readers (Benjamin and Schwanenflugel, 2010), and that children's ability to use aspects of lexical prosody is predictive of reading fluency ability (Schwanenflugel and Benjamin, 2017). There is a lot more to explore in this area of prosody research, especially through the work of Schwanenflugel and her colleagues, but it is not relevant for the current project as it is.

Though the terminology can be confusing, it would seem reasonable to assume that prosody is an important aspect of reading well. Prosody can help reading by giving children the ability to recognise and use appropriate changes in intonation, whilst paying attention to the syntax and punctuation in the text (Kuhn et al., 2010). Prosody can aid phrasing of the text, chunking the written words into meaningful units for understanding by the listener, along with appropriate pauses and syntactic parsing to ensure meaning is conveyed clearly. This relationship needs to be explored through the developmental stages of reading, as though we may assume the appropriate use of prosody will transfer from spontaneous speech onto oral reading, we can't say for sure (Veenendaal et al., 2014). Rather, we need to measure speech prosody and oral reading in tandem, and disentangle its relationship with other important aspects of reading ability development (e.g. decoding) to try and figure out what makes prosody so special for fluency. For example, the importance that prosody skill could have for oral reading of multisyllabic words, to be able to produce the appropriate stress when reading these words in the context of the rest of the text could alter the meaning of the text if children don't know which syllable to stress.

2.8.2. Implicit Prosody

Through the preceding subsections I've focused upon the metalinguistic applications of prosody, of the sensitivity to aspects of language such as stress and intonation. This is clearly important for literacy, understanding how to segment language using this information and comprehend meaning from the speech stream through the choice of prosodic features. However, that still leaves the question of how one imposes these prosodic features onto written text. Unlike other alphabetic languages English does not have special notations to indicate these features, like stress accents in Spanish for example. Rather, prosody in the English language has to be inferred from the text, using prior knowledge. We know this occurs, as expert readers we are able to successfully garner meaning from written text with little problem, but how do we get here? To read silently it may be that an acute awareness of prosody is important.

This calls to mind the idea of the inner voice, which has been discussed much in the literature, stemming from a quote by Huey (1908) thinking about the phonology and pedagogy of reading ability, saying that "the inner saying or hearing of what is read seems to be the core of ordinary reading, the "thing in itself", so far as there is such a part of such a complex process". The issue with talking about an 'inner voice' is attempting to provide a concrete definition of what an inner voice is. Does one develop an inner voice early on through conversation, or does it develop through learning the prosodic cues of the language? It seems reasonable that the population at large recognises that they have this voice when reading, but in attempting to measure and quantify that it becomes more complicated. It is only recently that technology and ideas about reading have reached a point where these questions can be explored with some reasonable level of inference.

A major theory that continues to be discussed in the literature regarding this intersection of the inner voice and prosody is the Implicit Prosody Hypothesis (hereafter referred to as IPH). The IPH was posited by Fodor (2002), though it is important to note it developed from earlier discussions of these ideas by herself (Fodor, 1998) and Bader (Bader, 1998) a few years earlier. The basic concept of the IPH is that when reading silently, readers project a default prosodic contour onto a stimulus that is appropriate to the content. This projection may, in turn, affect syntactic ambiguity resolution, as the prosodic contour implemented by the reader could change the phrasing of the stimuli and affect understanding. This can be seen in sentences where the syntactic boundary is affected by the placement of a boundary, such as these examples from Fodor (1998); "The divorced bishop's daughter" / "The recently divorced bishop's daughter" / "The recently divorced bishop's daughter-in-law".

Fodor (1998) reflected upon the inner voice, arguing for the presence of a prosodic parser which is operating in parallel to the more well understood syntactic parser in sentence reading. The syntactic parser is, as it sounds, a function of reading ability that allows readers to utilise their knowledge of syntax to comprehend the meaning of what they are reading. This may be calling on previous knowledge from previous sentences to aid current understanding, understanding the context of the sentence within the current piece being read, and how the construction of the sentence can help to guide the reader to correct understanding. Fodor argued that this was true, but that the prosodic parser could supersede, or at least inform, this correct understanding also. When we read, Fodor said, we prefer sentence material to split into phrases of roughly equal length, as this aids our ability to process meaning. As such, when presented with a longer sentence, the likelihood of chunking the sentence into smaller phrases is more likely. Depending upon how the chunking occurs, it can have implications upon syntactic attachments. Fodor named this prosodic parser outcome the *same-size sister constraint*, that the prosodic parser prefers balanced phrases, and this can alter comprehension, depending upon how the sentence is balanced. Following this logic, we may assume that readers are trying to find the most natural prosodic contour for what they are reading, and through implicit balancing of the sentences comprehension is manifested.

One of the issues that has continued to plague implicit prosody research is how to measure such a phenomenon. The idea itself sounds very much like a paradox, the idea that inner speech generates prosodic information for understanding that is not provided by the written text. as if the readers impose a prosodic contour on the text to help themselves (Ashby and Clifton, 2005), it is hard to conceptualise quantifying such a skill. Fodor (2002) herself reflected on the difficulty of measuring implicit prosody, which often involves the manipulation of prosody, which in turn can end up manipulating another linguistic feature, making the methods difficult to single out implicit prosody skill. An example of a method used in the literature is that of syntactic ambiguity resolution. This is alike to garden-path sentences, where a sentence or phrase is presented which has been designed to lead the reader to some ambiguity in its reading, maybe through the manipulation of grammar or in the structure of the words chosen. One classic example is "The old man the boat", with the design of the sentence leading the reader to assume the phrase is about the "old man", but readers then realise this doesn't make sense and re-read the sentence wherein it is "the old" who "man the boat". This manipulation highlights the assumptions readers make while reading a text, and that when our syntactic expectations are not met, or a close reading is not being practised, then incomprehension or re-reading occurs.

A method to investigate this was designed by Trueswell et al. (1999), to explore how children read sentences using eye-tracking. Children were presented with a table of objects and given instructions to perform an action. The action was either unambiguous ("put the frog that's on the napkin in the box") or ambiguous ("put the frog on the napkin in the box"), with the verb "put" always the action to be performed. The ambiguous sentences were always designed to be syntactically tricky, as in the example above there were two plush frogs, one just a frog and the other a frog sitting on a napkin. The novel aspect of this study was the use of eye-tracking technology, which with the development of easier-to-use systems has become a more viable measure of reading ability with children. In this study, Trueswell and colleagues used a visual scene with the objects on the table, and measured where children were looking in the scene when they heard the sentences, allowing the researchers an implicit measure of how the children comprehended the ambiguous sentences based on the objects they looked at.

The relationship between implicit prosody and reading continues to be explored in the reading science field. Understanding how the reader imbues written language with prosodic cues akin to those present in spoken language could be an important aspect of understanding reading comprehension. With the wider availability of technology such as eye-tracking, which provides an implicit window into the effect prosody has on reading ability. Further discussion of this aspect of prosody is beyond the scope of this project, but I regard it as an important aspect of further research, and discuss this more in Chapter 5.3.2.

Summary

The two prosodic skills discussed above are included in the current project as reflections of the complexity that occurs when trying to explore prosody's relationship with reading. In prosodic reading we have the discussion of how prosody may aid oral reading ability, and provide a means for self-teaching and to read with a natural voice. Then with implicit prosody, with a phenomenon that research which sounds logical and which research has found success in exploring, but to measure is quite difficult and relies on inference from other tasks. It is important to note, that from this author's point of view that both of these aspects of prosody are still fertile grounds for further research, especially implicit prosody which with more research could be solidified in its importance for early silent reading development. But, they reinforce the point that prosody as a concept is hard to pin down, and measuring it successfully is even harder. The research discussed in Section 2.4 and 2.6 focused mainly on behavioural measures of prosody, whether it is the DEEdee task's stress sensitivity, or the PEPS-C 2015 which attempts to cover a wide variety of prosody skills within a simplistic behavioural tests format, in the prosody and reading literature a lot of the measures used are binary correct-incorrect quantitative tests. It may be to get the full scope of prosody's influence upon reading we need to approach prosody from various angles, and see if the end result changes with these interactions. Unfortunately this was beyond the scope of the current project, but at the end of the thesis I attempt to discuss this in Chapter 5.3.2.

2.9. Prosody, Reading and the Current Project

Throughout this sub-chapter I have explored prosody and what we know currently about its relationship to reading development. I discussed the varying definitions of what prosody *is*, and how this can affect what measure a researcher will choose to represent prosody. The definition being followed for this project is of prosody as a linguistic aspect of spoken language, and divide this into receptive and expressive skills to aid understanding and comprehension. I did not explore prosody with any speech analysis measures, nor did I design a study to manipulate prosody in the tasks. I was interested in using a measure to look at children's competence on measures that attempted to single certain prosodic traits for measurement. For the current study I chose to use the PEPS-C 2015 battery measure, as it contains multiple behavioural tests of prosody, using both non-verbal (receptive) and verbal (expressive) tasks, an attempt to measure both understanding and use of prosody. I was able to create composite variables of receptive and expressive prosody, which I thought was an interesting approach to compare which different aspects of reading ability within the SVR more strongly related to receptive or expressive prosody ability.

There were many other directions that could have been taken to operationalise prosody in the project. I chose to reflect on these in discussing prosodic reading and implicit prosody, and believe that these are avenues of research that could be brought together to get a breadth of knowledge about prosody as a concept in relation to reading. However, for the moment this project has reduced this to a focus on a definition of prosody which differentiates the skill into a receptive and expressive form, which was hoped to bring out interesting observations when the project was completed.

Furthermore, I did not include a measure of morphological awareness in the current project, which as discussed above may be an important indirect mediator between prosody and reading comprehension. This was partly due to wanting to focus on the SVR components in relation to prosody, but future studies should include morphological awareness measurements as it seems from the projects highlighted that this may be a major aspect of later reading ability, and the relationship between morphology and prosody an interesting avenue of potential research.

Chapter 3. Methodology

The current project explored the relationship between children's prosodic abilities and reading comprehension, controlled by measures representing the components of The Simple View of Reading. Furthermore, two other tasks were included as control measures for the effects of prosody that were not part of the SVR. These were a phonological awareness task and a non-linguistic verbal working memory task. Each of the measures is discussed in more detail below. Following this, a summary of the participant criteria used for selection is discussed. The sample of children were all taken from three Year 3 (ages 7-8) classrooms from a single school in the North East of England. Arrangements were made between the lead researcher and the school through an intermediary group called Research Buddies, who set up the initial relationship between the groups. Next, a discussion of the procedure for collecting the data is outlined. These were the steps the researcher collecting the data followed with each participant to ensure they conducted the measures the same. Finally, a summary of the statistical procedure that will be used to answer the research questions. The statistical measures decided upon were descriptive statistics, correlations, regression analysis and mediation analysis. The reasons why these measures were chosen will be discussed, and the procedure to look at the data will be made clear.

3.1. Current Study

Despite the importance of understanding reading comprehension in a child's literacy development, only a few studies in the literature have investigated the contribution of prosody to reading comprehension as the major focus of their studies (Chan et al., 2020; Clin et al., 2009; Deacon et al., 2018; Holliman et al., 2010b; Kim and Petscher, 2016; Lochrin et al., 2015; Whalley and Hansen, 2006; Whalley, 2017). To investigate the relationship further, the current project utilised the PEPS-C (Peppé, 2015) as used in Lochrin et al. (2015) to investigate prosody and reading comprehension. The choice was made to create two composite measures of prosody - receptive prosody and expressive prosody. The hope was that I could investigate the difference in contribution between the two aspects. This was measured in tandem with measures reflecting the components of The Simple View of Reading framework, as control and outcome measures, as well as an overall framework of reading ability. Below are the research questions and hypotheses for the current thesis.

One of the unique contributions the project sought to explore was the individual contributions the receptive and expressive prosody measures had to reading comprehension. This was a gap that I identified when discussing the various projects which looked at prosody and reading comprehension, with only Lochrin et al. (2015) making explicit reference to this difference. I was interested in whether either receptive or expressive prosody would have a stronger relationship with the reading comprehension measures we used. Based on (Lochrin et al., 2015)'s project, I predicted the receptive prosody measures would be more strongly related to the word reading measures in this project, and the expressive prosody measures to the linguistic and reading comprehension measures.

The project had two main research questions to explore:

3.1.1. Research Question 1: Does prosody provide a unique contribution to children's reading comprehension ability, over and above the other predictors?

The major question I had was whether prosody was a direct contributor to reading comprehension ability, or if the components of the Simple View of Reading affect this relationship. Based on the literature review the framework of the SVR has a strong research base, suggesting decoding and linguistic comprehension have are necessary predictors of reading comprehension ability. This project wants to investigate if prosody is predictive of reading comprehension when entered into the SVR framework.

Hypothesis One: The measures of prosody in our project will have a significant relationship with the measures of reading comprehension ability.
This hypothesis is based upon the results of four previous studies of prosody and reading comprehension (Holliman et al., 2010b; Lochrin et al., 2015; Whalley and Hansen, 2006; Whalley, 2017) in which prosody showed a direct relationship with the measure of reading comprehension in the respective relationship. The expectation was that the receptive and expressive aspects of the PEPS-C 2015 prosody measure would have a small but significant direct effect with the measures of reading comprehension in this study.

Hypothesis Two: The measures of prosody in our project will have a significant relationship with the measures of reading comprehension ability, when controlling for the independent variables of the The Simple View of Reading framework
To expand on the previous hypothesis, I wanted to see if prosody would be predictive of reading comprehension when controlling for the SVR variables. The prediction was that prosody would still be predictive of reading comprehension, but due to the well-established relationship between decoding and linguistic comprehension to reading comprehension, its contribution may be small but significant.

3.1.2. Research Question 2: Which one of receptive or expressive prosody is a more consistent predictor of reading comprehension?

The main gap in this project I investigated was the contribution of receptive and expressive prosody to reading comprehension. This was inspired by Lochrin et al. (2015), but I took this study further by creating separate composite measures of receptive and expressive prosody, to investigate their individual contributions. As mentioned above, I expected the receptive prosody measures to be more predictive of the word reading measures (TOWRE-2 Short Word Reading, TOWRE-2 Phonemic Decoding Efficiency, YARC Short Word Reading Test) in the project, and the expressive prosody tasks for the linguistic and reading comprehension measures. This is due to expressive prosody requiring an individual to use their speech ability to use prosody out loud, which requires processing and understanding where to place stress, intonation and timing to ensure a word, phrase or sentence has the intended meaning. Taking this knowledge into consideration suggested expressive prosody would be the most consistent predictor of reading comprehension in this project.

3.2. Measures

Listed below are the measures I used for the current project. Each measure of prosody, decoding, listening comprehension, reading comprehension and control measure was chosen because of its use in prior projects investigating the speech and reading relationship. All measures were quantitative, and required children to respond verbally or with non-verbal actions to provide their answers.

3.2.1. Prosody

Deciding how to measure prosody was an important decision for the project. Many measures of prosody were discussed in Section 2 for consideration to use in the current study. The decision was made to use the PEPS-C 2015, an updated version of the old PEPS-C battery test created by Peppé and McCann (2003).

PEPS-C 2015

The Profiling Elements of Prosody in Speech-Communication is a collection of tasks to investigate prosodic ability with children and young people created by Peppé and McCann (2003). The task was devised as a research and clinical tool for use with autistic children, as a way to measure their ability to express and understand different aspects of prosody and attempt to specifically highlight which of these aspects individual children were weaker at understanding or using. The original test was made up of 16 tasks, measuring different specific prosodic abilities (e.g. stress, affect, boundary awareness) at different linguistic levels of language (word-level, phrase-level, sentence-level). Gibbon and Smyth (2013) attest that the strength of the PEPS-C is that it takes a psycholinguistic approach to measuring prosodic understanding and use, attempting to identify likely issues in autistic children's communication skills.

The current version of the PEPS-C is PEPS-C 2015, a revised edition combining elements of the older Research and Clinical variations of the tasks. This current edition has 14 tasks addressing receptive and expressive prosodic skills in parallel tasks matched by the stimuli used. The tasks take 40-60 minutes to complete and are conducted using a mixture of verbal response and non-verbal computer interactions. The receptive prosody tasks are simple to implement, requiring children to listen to stimuli and respond by selecting stimuli on the computer. The expressive prosody tasks, on the other hand, require the researcher to record children's responses via a specially defined keypad, provided by Dr Peppé (see Fig 3.1). The PEPS-C 2015 is a semi-automated battery test, in that the recording of answers and playing of stimuli was programmed automatically by the software, the only input the researcher and participant need to provide is a click or a key press for the task being conducted. Following this, the output of the scores is calculated and displayed automatically by the software through a pre-programmed Excel spreadsheet.

Before beginning the tasks, the researcher conducted a familiarisation task with the children. This was to ensure the children understood certain concepts that were to be used in the tasks to come, such as understanding stress differences between words and stress boundaries. Appendix B.3.1 provides screenshots of the familiarisation tasks completed by the participants. This took between five and ten minutes to complete, and gave the researcher ample chance to help and correct children's interpretation of the stimuli.

To record the verbal responses I used the keyboard provided by Dr Peppé with the PEPS-C 2015 materials. The keypad was designed to work with the PEPS-C spreadsheet without any programming on the researchers part, meaning they could just plug it into a USB 2.0 port and it would work. Furthermore, the researcher recorded the child's voice as they completed each of the expressive tasks to revisit and check their responses. Each task was scored out of a total of 16, further divided into expressive and receptive results out of a total of 112, and the total score for all tasks together was out of 224. Appendix B.3 describes which buttons were pressed in each task to record certain responses.

As discussed earlier in this section, the PEPS-C 2015 tasks were designed so that there was parallel receptive and expressive tasks for each prosodic skill being assessed - seven receptive tasks and seven expressive tasks. These tasks were presented in an alternating order (receptive then expressive, then receptive etc.) and the same overall order was conducted with each student.

Here we provide two examples of a receptive and expressive task, focusing on the Turn-End tasks. This task investigates the prosodic function of using speech to indicate whether the utterance is a question or statement. For this task single word utterances were used e.g. 'Carrots?' and 'Carrots'. For the receptive version of the task, children saw two pictures appear on screen (as seen in Figure 3.2) which represented either a question (the left picture) or a

Figure 3.1 PEPS-C Keypad: Used by the researchers to record the responses of the children in the PEPS-C 2015 tasks.

statement (right picture). The children then heard a pre-recorded voice say a single word as a question or a statement, and had to decide which picture best represented what they had just heard. In contrast, the expressive version of the task (seen in Figure 3.3) displays just a single picture, either the question or statement version. The child is then asked to say the food item on screen with the appropriate intonation, denoted by the picture being displayed. The response is then rated by the researcher identifying if the appropriate form of vocal intonation was used. This example is a good representation of how the tasks were presented. Appendix B.3 provides examples of the remaining tasks with longer descriptions and screenshots. Furthermore, see Table 3.1 for a summary of the fourteen PEPS-C 2015 tasks. They were completed in the order presented in the table.

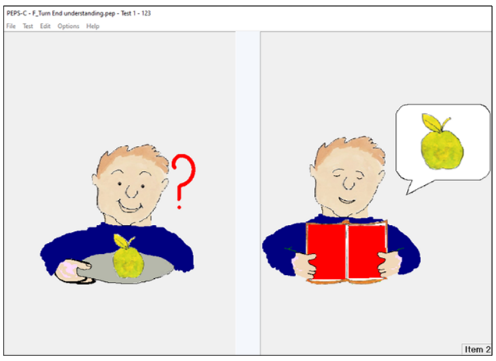

Figure 3.2 PEPS-C 2015 Receptive Turn-end Task

Figure 3.3 PEPS-C 2015 Expressive Turn-End Task

PEPS-C 2015 Task Name	Description
Receptive Short Item Discrimination	Tested children's ability to perceive intonation differences in short utterances.
Expressive Short Item Imitation	Tested children's ability to listen to and imitate different forms of prosodic intonation in short utterances.
Receptive Turn-end	Tested children's ability to perceive the difference between questioning and declarative intonation.
Expressive Turn-end	Tested children's ability to produce appropriate intonation (declarative or questioning) in response to a visual stimulus.
Receptive Affect	Tested children's perception of intonation to indicate like or dislike.
Expressive Affect	Tested children's ability to use intonation to denote like or dislike.
Receptive Lexical Stress	Tested children's ability to perceive stress placement in two-syllable words.
Expressive Lexical Stress	Tested children's ability to produce appropriate stress placement with two-syllable words.
Receptive Phrase Stress	Tested children's ability to comprehend differences in phrase stress.
Expressive Phrase Stress	Tested children's ability to produce the appropriate phrase stress for the word/s provided.
Receptive Boundary	Tested children's ability to disambiguate phrases that were syntactically ambiguous.
Expressive Boundary	Tested children's ability to produce the appropriate pause boundary based on stimuli presented on screen.
Receptive Contrastive Stress	Tested children's ability to perceive stress difference in a phrase.
Expressive Contrastive Stress	Tested children's ability to produce appropriate stress in response to an incongruous statement.

Table 3.1 PEPS-C 2015 Task List: Short summaries of the tasks implemented with the children for the current project. For more detailed descriptions of each task see Appendix B.3.

3.2.2. Decoding

Three measures of word reading ability were used in the current project. The three tasks were all variations on a word recognition task; two were measures of word reading fluency, and one a more traditional word reading task. These measures were chosen to capture the children's ability to decode words. Each have been used in previous studies, but the only ‘pure’ measure of decoding in this selection is the TOWRE-2 Phonemic Decoding Task. This is because the other measures used capture decoding skills, but the tasks can be completed through other means (e.g. sight word reading). The phonemic decoding task however uses nonwords, words the children should never have encountered before, so their internalised knowledge of decoding strategies learnt through phonics should come into use.

TOWRE-2: Test of Word Reading Efficiency - Second Edition

The first measures of decoding ability were taken from the Test of Word Reading Efficiency - Second Edition (hereafter referred to as TOWRE-2) created by Torgesen et al. (2012), a measure of an individual’s ability to correctly pronounce and identify as many words and phonemically irregular words within a time limit. More than just being able to recognise words, children were challenged to read the lists, testing the fluency of their word reading ability. Children were not expected to read every word on the list, the purpose of the task is to see how quickly and fluently children could decode the words on the page correctly. This reflects not just that children recognised the words and decoded them correctly, but it also showed an automaticity to their decoding, where they were not solely relying on sounding out each word phoneme-by-phoneme.

TOWRE-2: Sight Word Efficiency: Children were presented with a list of 108 words, which increased in difficulty the further children made it through the list. Children had 45 seconds to read as many words as they could. After 45 seconds children were told to stop and the researcher made a note of the last word read.

TOWRE-2: Phonemic Decoding Efficiency: This task followed the same procedure as the Sight Word Efficiency task, except the stimuli were not real words. Examples of the nonwords children had to read were ‘ba’, ‘nasp’ and ‘luddy’, each made up of real phonemes, but combined in a way to make a pronounceable word that does not exist (see Appendix B.4.2 for the full list). This meant the children could not rely on sight-reading as in the Sight Word Efficiency task, the nonwords had to be decoded when encountered. There were 66 nonwords in total the children could read, and they again had 45 seconds to read as many of the nonwords as possible.

YARC Short Word Reading Test

The YARC Short Word Reading Test is a word-reading subtest of the YARC (York Assessment of Reading for Comprehension) created by Snowling et al. (2011), a word-reading task where children were instructed to read as many words as they could. Children were directed to read a list of 60 words which steadily progressed in difficulty. Unlike the previous decoding tasks, the YARC SWR was not timed, allowing children more time to decode the words than the previous tasks. See Appendix B.6.1 for the full list of words.

3.2.3. Linguistic Comprehension

As discussed in the literature review, linguistic comprehension is a definition that is wide-ranging and hard to pin down. The task I chose required children to listen to speech and respond appropriately (in this case, choosing the correct illustration from a choice of four), which can be defined as a listening comprehension task based upon the method being used to assess comprehension, while also falling under the umbrella term of linguistic comprehension.

TROG-2 Listening Comprehension - Modified

The measure of linguistic comprehension in this project was the Test for Reception of Grammar Version 2 (TROG-2) created by Bishop (2003). The TROG-2 is a receptive language test where children listen to the researcher read sentences with differing types of grammatical complexity, then choose from four pictures which they feel represents the sentence they just heard (to see a visual example of this task see Appendix B.5.1). The researcher read the sentence to the child (e.g. "The book that is red is on the pencil") and the child would provide their response. Of the 80 items in the TROG-2, I used 40 of the sentences for measurement. This is because I created a modified version of the TROG-2 for reading comprehension as Whalley (2017) did, as a matched measure. This meant choosing only the sentences where I could reverse the nouns for the sentences, and the sentences read still made sense and were had a picture associated with this meaning.

3.2.4. Reading Comprehension

There were two measures of children's reading comprehension ability. The first was the modified version of the TROG-2 task mentioned above, altered to match the stimuli above whilst focused on children's reading comprehension ability. The second, the YARC reading comprehension test, is a more traditional standardised measure of reading comprehension, which required children to read two extended passages of text and answer comprehension questions following this.

TROG-2 Reading Comprehension - Modified

The TROG-2 Reading Comprehension task was a modified version of the original TROG-2 task modified for reading comprehension. This was done by reversing the noun placement in sentences that could be noun-reversed and were still reasonably comprehensible and had a corresponding picture in the test materials (e.g. 'The **duck** is bigger than the **ball**' → 'The **ball** is bigger than the **duck**') to create a balanced set of 40 sentences between the two measures. This allowed us to use the stimuli twice for different purposes, and I counterbalanced each participant in turn so I could balance for priming of the sentences for the tasks.

The instructions and method for the TROG-2 Reading Comprehension task did not differ from the listening comprehension version, except that children read the sentences instead of listening to the researcher say them. The sentences were presented in 18pt Calibri font at the top of an A4 page with the four pictures displayed below. Children were instructed to read the sentence aloud to the researcher, to ensure the researcher knew the child was reading the full sentence. Children would then point to the picture they thought the sentence was referring to, and the researcher recorded their answer. An example of this layout is presented in Appendix B.5.2.

YARC Reading Comprehension Test

The YARC reading comprehension test involved children reading two extended passages of text (one fiction, one non-fiction) aloud to the researcher, followed by eight comprehension questions about each text, sixteen questions in total. The YARC manual provided suggestions for the text choices depending on year group based upon standardised data collected by the authors. For Year 3 children, the recommendation was to use texts levels 3 and 4 for reading from the booklet. Scans of these texts can be seen in Appendix B.6.2.

The task began with the researcher presenting the child with a passage of text to read, and asking them to read it as accurately as possible and to request help if they did not know how to say a word. As the child read, the researcher kept track of any errors the child made and would correct them. This was to ensure children understood the full context of the passage they were reading, as an incorrectly read phrase or word could alter the meaning of the text. Also, the researcher recorded how long it took to complete the passage with a stopwatch.

Immediately following the reading of the passage the researcher asked the child eight comprehension questions about what they had just read. The child could look back at the text if they needed to, to help them answer the questions. Each question had specifically worded answers provided by the YARC manual, but some answers which were vague and the researcher could probe for a more specific answer. See Appendix B.6.2 for the list of questions and example answer sheets. These same instructions were repeated for both texts, the only difference being that the second text was longer than the first.

3.2.5. Other Measures

This final subsection includes the two tasks that have been included in an attempt to control for two factors that contribute to reading comprehension ability. The first is a measure of phonological awareness, an aspect of language that has been well-established as a necessary skill for early readers. The second task is a measure of verbal working memory, specifically short-term working memory with numbers. It is important to note that there are many other non-linguistic skills that have been investigated in the literature which contribute to making a successful developing reader, examples including attention monitoring, visuospatial skills, and non-verbal working memory which can aid general reading ability. Controlling for these skills does not begin to account for all the possible factors which contribute to reading ability, but they were included to capture some of the unique variance in this project.

CTOPP-2 Elision

The measure of phonological awareness chosen for the project came from the Comprehensive Test of Phonological Processing - Second Edition (CTOPP-2) by Wagner et al. (2013). The CTOPP-2 is a battery test of different measures for measuring phonological awareness, phonological memory and rapid naming ability with children and young adults. The task chosen for this project was the elision subtest of phonological awareness. This task required children to listen to the researcher read scripted sentences such as:

“Say CUP...(child responds)...Good! Now say CUP without saying the /k/"

In the above example, the answer would be “UP". The 34 items got increasingly harder as the children got further down the list, such as being asked to remove phonemes from the middle of words:

“Say PIXEL..(child responds)...Good! Now say PIXEL without saying the /s/"

The task was stopped if children got three items in a row incorrect. The elision task was chosen as the representative task for phonological awareness as there is a low probability for ceiling effects because the stimuli at the end of the task are very difficult, even for adults. This consideration was made because phonics teaching is such a major component of teaching in UK schools, that by Year 3 children have been using phonics techniques to help them read for two-to-three years, so to have a sufficiently difficult task was a necessity. The full list of words can be seen in Appendix B.1.

Working Memory Test Battery for Children (WMTB-C) - Backward Digit Recall

Taken from the WMTB-C by Pickering and Gathercole (2001) the backward digit recall is a measure of short-term verbal working memory. Children were instructed to listen to the researcher recite a list of numbers, and repeat the numbers back in reverse. For example, if the researcher read the numbers "2, 4, 7", the child should respond with "7, 4, 2". The list of numbers got longer as the task went on, starting with two numbers to recall, and ending with seven numbers if the child got far enough. The numbers were presented in blocks of six, meaning there were a total of 36 possible lists the children could recall. The numbers were read by the researcher with a monotone inflection and a slow speed to ensure each trial was read as similarly as possible. Once the child correctly recalled four of the six lists correctly they moved onto to the next section. The task was stopped if children made three or more errors within a block. The list of numbers the researcher read to the children can be seen in Appendix B.2.

3.3. Equipment

To be able to collect all this data there was a large amount of equipment required to ensure it went successfully. All of the data was collected using a Windows laptop, to ensure portability when going to schools, and so data entry could be done remotely during breaks in collection. The laptop was selected to ensure it could run all the software required for the research.

To present the auditory stimuli for the prosody tasks I used a pair of Sennheiser HD 25 Basic Edition headphones. They are a high-end model of headphone that provides a high frequency response (16-22000Hz), suitable for presenting the audio stimuli at high quality. They were also chosen for the smaller than usual ear cups, which fit well for child participants and can be adjusted for particular head sizes in a simple manner.

I recorded the children for certain tasks using a Tascam DR-05, a small portable dictaphone that recorded high quality speech in a WAV format. I only used the recordings for the lead researcher to check their data entry was correct. I discuss in Section 5.3.1 how these recordings could have been utilised for the research more effectively.

Physical stimuli was also brought that could not be used via the computer. This included the booklets for the YARC Reading Comprehension Passage Reading Task, the decoding task word lists, and the pictures for the TROG-2 tasks. All other tasks were conducted by the researcher via the laptop.

3.4. Participants

The final sample for the project were 51 Year Three children from a single school in the North East of England. This particular age group was chosen to work with based upon the Government's own curriculum (Department for Education, 2013), as the curriculum states that Year Three is when reading comprehension teaching becomes a focus of English lessons. However, this does not mean that reading comprehension teaching had not begun prior to children coming into Year Three, but this knowledge provided us with a specific time frame in which to select participants. I believed this was a good time point to focus on, as it represented the point where an explicit transition was being made towards independent comprehension of written text by the children.

The children were recruited through cooperation with the Newcastle University scheme Research Buddies (more information available at https://bit.ly/2Jskmhf). This scheme was devised by members of staff within Newcastle University who were interested in creating a partnership between the university and local schools in the area, to create a network of schools who would be willing to accept researchers to collect data in their school.

Table 3.2 presents an overview of the demographic group I was working with for the project. I focused on working with a typically developing group of children, which in this case meant children without neurological or learning difficulties as defined by their parents in the questionnaire I shared. The children's neurological or learning difficulties were identified by the parents in the questionnaire I completed. I worked with all children who responded to the request for participants, as I did not want to leave any child out. In the end, this meant excluding five children who had prior neurological or learning difficulties from the final data set (one child with profound deafness, two children with ASD, one child communication difficulties, and one child just identified by their parent to have learning difficulties).

Beginning with the age in months our participants spanned the entire range of the Year 3 ages, with a 12-month range of ages from 93 to 105 months. Next, the sex variable, in which I have almost got a 50-50 split of female and male participants (57%/43% respectively) from our data collection. This 50-50 split would likely have been achieved if participants had not needed to be excluded for the reasons above. Lastly, handedness, and I have a split of almost 90% right-handed students.

I did not use this demographic information any more through the rest of the study for analysis, as the group sizes for each would be too small. Thus, from here on out I conducted all statistical analysis for the Year 3 class as a whole.

Participant Characteristics (N=51)	N	Mean	Range
Age in Months		99 (SD = 3.6)	93 - 105
Sex			
Female	29 (57%)		
Male	22 (43%)		
Handedness			
Left	6 (12%)		
Right	45 (88%)		

Table 3.2 Descriptive Statistics for Participant Information

3.5. Procedure

The headteacher consented to the research taking place in the school, and agreed to distribute consent forms and other documents to the parents of the children in the age group to be worked with. These documents were as follows:

- Parental Consent Form (see Appendix A.1) - a consent form detailing what the study was about and the ethical procedures the project would follow. Parents had to sign and send back this consent for their child to be able to take part.
- Parental Information Sheet (see Appendix A.2) - a short two-page document detailing and answering expected questions about the project's aims and procedure.

- Parental Questionnaire (see Appendix A.3) - a short five-question questionnaire asking parents for more information about their child. This was used also to identify any children with learning or neurological difficulties who may find the tasks particularly difficult.

Once the researcher had received the consent forms from their parents, those who identified they wished to take part were taken to a quiet space to complete the tasks. The researcher discussed when the children could be taken out of class, and took each child out of the classroom one-by-one to complete the tasks. Children were seated and the researcher presented them with a consent form, reading the text on the consent to the child to ensure they understood what they were agreeing to (see Appendix A.5 for a scan of the child consent form). Furthermore, it was important for the researcher working with the child to highlight that they could stop the tasks at any time and return to their classroom without questions being asked. All of the children agreed with what was said on the consent form and took part in the tasks.

The researcher worked with the children over three individual sessions of 30-40 minutes each, so as not to fatigue the children and also to ensure they did not miss significant amounts of teaching in the classroom. Tasks were undertaken in the same order for each child, except for the TROG-2 listening and reading comprehension tasks which were administered in a counterbalanced order for every other child. The order was as so:

- **Session 1** - TOWRE-2 Sight Word Efficiency, TOWRE-2 Phonemic Decoding Efficiency, YARC Short Word Reading Test, CTOPP-2 Elision and TROG-2 Listening Comprehension/Reading Comprehension.
- **Session 2** - PEPS-C 2015 (Pre-test vocabulary check and first 8 tasks) and TROG-2 Listening Comprehension/Reading Comprehension.
- **Session 3** - PEPS-C 2015 (The final 6 tasks), Backward Digit Recall, and YARC Reading Comprehension.

The total time to complete all of the tasks was on average 2 hours, depending upon how motivated children were on the day of and how difficult they found individual tasks. At the end of the tasks, each child was given a certificate to celebrate their role as a ‘junior researcher’ to thank them for taking part.

3.5.1. Data Management

Research continues to become ever more concerned with digital preservation, meaning digital data security is a more pressing concern than ever. Related to this study, I was collecting data from young children, a group whom have little autonomous control over data and as such need to be treated with respect and security.

Metadata

To begin with, each child was given a unique number to identify them, which was only meaningful to the lead researcher. This metadata ensured the researcher could organise the data meaningfully, but ensure anonymity for each of the children. This code was the school class number worked with (01, 02 or 03 - no relation to actual class number, just whichever was worked with first), and the order of the participant worked with in the class. For example, if a child was in the second class and the fifth child the researcher worked with, their code was 0205. This prevented any identifiable information being available from the file names.

Data Access and Sharing

The only individuals with access to the raw data was the lead researcher. Before sharing any data with other individuals the data was processed and collated into a .csv file with no specific identifying information beyond the participant number. The data was shared with the lead researchers supervisors for discussion and checks that analysis was suitable, but before the supervisors were privy to the data the lead researcher ensured the data was:

- Anonymised: this was done to begin with by the metadata procedure used to ensure the researcher alone could identify the children if needed to check consent or remove data if requested, but no-one else would be able to identify who the data was referring to. Furthermore, no identifiable data was included in the data output, such as the learning or neurological difficulty the children had as this could be potentially identifying if an individual figured out the school the lead researcher worked with.
- Consented for sharing with the supervisory team: as part of the consent process the parents/guardians of the children were informed that the children's data would be shared for discussion with the supervisory team, but that this data would be anonymised prior to sharing.

3.6. Data Analysis Procedure

3.6.1. Statistical Procedure

After collecting, organising and checking the data the next step was to analyse and interpret the output. Data analysis was conducted with R (R Core Team, 2019), an open-source and widely used statistical programming software used much more in recent years in psychological and linguistic research. The choice to use R was motivated by the trends for open science in the field, allowing for the sharing of R scripts online for others to run the code and see the same outputs I ran, helping to make the project reproducible. Furthermore, it allows others to challenge the assumptions of the output more effectively if they find errors in my code, or provide an alternative coding strategy to analyse the current data. It allows others to take this data set and explore it for themselves, which could be useful for future studies. Finally, conducting the data organisation in R allowed us to track the stages I went through when cleaning the data, for evidencing choices made and tracking decisions if forgotten. This luxury is not there in SAS or SPSS, and this led to me ultimately choosing R for this reason.

The data collected was explored and analysed in four stages:

Step 1: Summary Statistics

First, I discussed the summary statistics for each variable analysed discussed in this chapter. These were descriptive statistics for discussion and ideas of the variance and dispersion of the data. For the continuous variables (the majority of the tasks), I output the mean, standard deviation, median and range of values. For the categorical variables (sex and handedness) I looked at the counts.

Step 2: Correlations

Second, I explored the correlations between each of the tasks. To conduct our correlations I first conducted assumptions tests to ensure the data were suitable for using a parametric version of the correlation test. I conducted the correlation with and without the transformed variables I identified (CTOPP-2 Elision and TROG-2 Reading Comprehension (Modified)) and found no significant difference in correlation output, only minor fluctuations. As such, the correlogram seen in Figure 4.1 was conducted with the data in its original form.

The correlation was conducted using the corr.test function in the R *psych* package (Revelle, 2019), which allowed us to select the method of correlation I wanted to use, in this case I used the parametric Pearson correlation coefficient for our testing. Furthermore, it allowed us to select the p-value adjustment criteria I wished to use, to account for multiple correlation tests. For this I used the Benjamini-Hochberg p-value adjustment method. This controlled for the false discovery rate, i.e. trying to control for the amount of significant results that are actually false positives. This test is less conservative than the other commonly used p-value adjustment method, the Bonferroni correction. I created a correlelogram to represent the multiple correlations measures conducted (see Figure 4.1 in Chapter 4.2). This was completed using the *corrplot* package created by Wei and Simko (2017), allowing us to display the correlation coefficients and the p-value to explore the correlatory relationships between the variables.

Step 3: Regressions

I created multiple regression models with various combinations of the task variables to provide answers for the various research questions and hypotheses. In particular, I wanted to look at the relationship between the Simple View of Reading measures and prosody i.e. looking at the predictive strength of prosody to reading comprehension after controlling for the measures of decoding and listening comprehension.

The first sets of models in Section 4.3.1 were designed to help us understand the relationship of the prosody measures as predictors of the reading comprehension and SVR variables, giving us an idea of prosody's contributory relationship to these outcomes.

The second set of multiple regression models in Subsection 4.3.2 were designed to help us answer our second research question, that of whether prosody still contributed to reading comprehension after the components of the SVR were added as controls. The two models included the totalled measures of expressive and receptive prosody, all the measures from the SVR (decoding and linguistic comprehension), and finally the control measures (CTOPP-2 Elision, and backwards digit span) as predictor measures. The two reading comprehension measures (YARC Passage Reading (Standard Scores) and TROG-2 Reading Comprehension) were the outcomes for these two models. For the TROG-2 Reading Comprehension model I conducted both an Ordinary Least Squares and Weighted Least Squares model, due to the heteroscedasticity present in the model. I present these regression models side by side in Table 4.7.

Furthermore, the predictors in these models were transformed to z-scores to allow for easier comparison of contributions to the outcome measures. As well, the decoding measures were combined to create a composite score for decoding. Both of these changes are discussed in Subsection 4.3.2.

Step 4: Mediation

Finally, I designed and computed four exploratory mediation models to look at the mediating relationship of the decoding and linguistic comprehension variables. This is discussed more in Section 4.4, but the mediation models are an exploratory model for suggestions of further study, as I completed a cross-sectional study I cannot determine causality in this instance. However, I felt in conjunction with the regression models I could make suggestions towards the relationship, and discuss this in reference to previous studies which used mediation analysis when reflecting on the results in the following discussion chapter. The four models are simple mediation models, consisting of a predictor, mediator and outcome. More detail regarding the mediation modelling procedure is described in Section 4.4, leading into the visualisation of the models.

3.6.2. Data Cleaning and Assumptions

An important aspect of the analysis procedure was cleaning the data for analysis. I created a separate R script to do some of this analysis, as well as doing the appropriate cleaning and assumption checks throughout the rest of the analyses. In this subsection I detail the assessments and methodology followed to clean the data and assess the assumptions. Any extraneous detail regarding this work can be found in the referenced appendices throughout the writing, and I refer to this aspects where appropriate in the Results chapter for understanding.

Outliers

Prior to beginning analysis I looked at the outliers in the data set. I wanted to identify which tasks had extreme outliers, and to make an informed and rationalised decision regarding whether to remove these participants from the analysis. To do this I used boxplot visualisations (see Appendix C.2 for these diagrams) and used the *diagnose_outlier* function from the dlookr package Ryu (2020). See Table 3.3 for the tasks where I identified the most extreme outliers to be. I used these diagnostics in conjunction to make decisions regarding altering the data for analysis.

	No. Outliers	Mean with Outliers	Mean without Outliers
Backwards Digit Span	1	14.39	14.16
TOWRE-2 SWE	4	60.96	62.74
TROG-2 RC (Modified)	4	30.71	31.19

Table 3.3 Outlier Summary: Table summarising the output from the *dlookr* function from the dlookr package (Ryu, 2020). The tasks identified above were the tasks identified to have the most extreme outliers for the tasks. Mean values for the tasks with and without the outliers are detailed.

I made the decision to not remove any outliers from the tasks initially. I tested the correlation measures with and without the outliers in a number of configurations, and there was no significant difference to the outcome. As such, I conducted these tasks with the outliers included. To see the longer discussion of these outliers, see Appendix C.4.1 for details. The main outcome, was that these were not erroneous scores, only children who scored at the extreme end of the scores collected. In our variables used I saw no obvious ceiling effects on the task, and as such it did not seem right to edit or transform the scores because some of the children performed very well. Thus, I kept the scores as they were.

It is important to note here, outliers related to the regression models will be discussed in the Results chapter. This is because these outliers are related to the residuals of the models, and as such taken on a model-by-model basis to decide on their inclusion.

Transformation of Data

An important consideration to be made when analysing the data was decisions regarding transformation of data. This is due to making corrections for the assumptions of normality and homoscedasticity that occur naturally through the measurement of data. In exploring the current data set, it was important to think about whether transformation was appropriate. Before conducting the correlations I analysed both visual and statistical methods of diagnosing these issues, to ensure I met the assumptions of the parametric correlation test as closely as possible. Individual discussions of transformation of the regression models are detailed in the appropriate sections of the Results chapter, as decisions regarding the assumptions of these models differed between the myriad models computed.

In this subsection I discuss the skewness, kurtosis and normality of the individual variables. I then discuss whether I did or did not transform certain variables in response to this information.

Skewness and Kurtosis

For the statistical aspect of testing whether I wished to transform the data I utilised statistical measures of Skewness, Kurtosis, and the Shapiro Wilks statistic for assessing normality. Furthermore, I transformed the output of the skewness and kurtosis measures using the *stat.desc* function from the pastecs R package created by Grosjean and Ibanez (2018). This function creates an output from our data containing the descriptive statistics discussed already, but also the extra measures of normality and dispersion seen in Table 3.4. One of the extra steps this function does is to produce an output for the skewness and kurtosis statistics divided by two standard errors. By transforming these values to z-scores I can use this information to guide us towards decisions regarding transformation. It means I can use statistics to see how likely the values of skew and kurtosis are to occur. By dividing the skew and kurtosis by the standard error twice I would say that values greater than one in this case are significantly skewed at a significance level of $p < .05$, values greater than 1.29 are significant at the $p < .01$ level, and above 1.65 significant at the $p < .001$ level. This guide for the scores was taken from Field et al. (2012), whom dedicated a section to discussing these particular statistics. It is important to note before discussing the results in 3.4 is that I did not take the statistical outputs from these tasks as complete truth for making decisions. The z-scores provide an extra guide towards making decisions regarding transformations of the data, and these decisions are made in conjunction with visual representations such as histograms and q-q plots. These visualisations can be seen in Appendix C.2.

If I take the output of 3.4 at face value, we see that the TROG-2 Reading Comprehension (Modified) measure is significantly skewed (-1.205 2SE) and a significantly leptokurtic kurtosis that is different from a normal distribution. This means the majority of the values are not represented around the mean, and that there was a wide spread in scores. This is likely due to the modification of the TROG-2 test, as it was designed as a listening comprehension task and not reading comprehension. Furthermore, with our age group comprehending the written word is more difficult from comprehending speech, so the spread of scores to create this distribution is not surprising.

The raw score for the YARC Passage Reading Comprehension variable was the other variable identified through the method above, with a significant platykurtic kurtosis (-1.030) according to the standard error method I detailed above. This is likely because the YARC Passage Reading

Score had a max score of 16, as there were 8 questions for each passage. This meant the spread of data was limited, compared with the standard scores which took into account the age of the children. For the final choice of variables I chose the standard score output as the representative variable for reading comprehension from the YARC, as it accounted for the age of the child and the YARC has been well-tested over time to account for these standard scores, and is one of the important aspects of it as a task.

	Skewness	Skewness (2SE)	Kurtosis	Kurtosis (2SE)
1	-0.630	-0.945	-0.846	-0.645
2	0.597	0.895	-0.001	-0.001
3	-0.062	-0.093	-0.555	-0.423
4	0.020	0.030	-0.824	-0.628
5	-0.657	-0.985	0.124	0.094
6	-0.019	-0.029	-0.324	-0.247
7	-0.422	-0.633	-0.901	-0.687
8	-0.804	-1.205*	1.361	1.037*
9	-0.345	-0.518	-0.744	-0.567
10	0.087	0.131	-1.351	-1.030*
11	0.245	0.367	-1.247	-0.951

Table 3.4 Table of statistics for the skew and kurtosis of the individual measures. The row numbers specify the specific task being referred to as follows: 1 = CTOPP-2 Total, 2 = Backwards Digit Span, 3 = PEPS-C 2015: Expressive Total, 4 = PEPS-C Receptive Total, 5 = TOWRE-2 SWE, 6 = TOWRE-2 PDE, 7 = TROG-2 Listening Comprehension, 8 = TROG-2 Reading Comprehension (Modified), 9 = YARC SWR, 10 = Passage Reading Comprehension (Raw Score), 11 = Passage Reading Comprehension (Standard Score). * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

Normality Assumptions

Skewness and kurtosis are aspects of the normality assumption, but there is also the Shapiro-Wilk statistical test, which is a measure of normality that compares the scores in the set with a normal distribution with the same mean and standard deviation, and assesses how different this distribution is from normal. This does have limitations, particularly in relation to the size of the sample, wherein larger samples are more likely to be significant from small deviations from normality, so the statistic needs to be used carefully. For our relatively small sample size the statistic should be reasonable. But, as discussed above with the skewness and kurtosis statistical output, this statistic was not taken as the absolute truth in representing normality. The statistic was used as a guide to make decisions regarding the assumptions of the correlation test.

The results of the Shapiro-Wilk test can be seen in Table 3.5, with six variables having been identified as significantly violating the normality assumption according to the statistical test; CTOPP-2 Elision, TOWRE-2 SWE, TROG-2 Listening Comprehension, TROG-2 Reading

	Shapiro-Wilk Statistic	Shapiro-Wilk P-Value
1	0.914	0.001***
2	0.967	0.171
3	0.976	0.391
4	0.975	0.353
5	0.952	0.038*
6	0.994	0.995
7	0.938	0.010**
8	0.931	0.006**
9	0.957	0.063
10	0.932	0.006**
11	0.938	0.010**

Table 3.5 Table of statistics for the Shapiro-Wilk statistic of the individual measures. The row numbers specify the specific task being referred to as follows: 1 = CTOPP-2 Total, 2 = Backwards Digit Span, 3 = PEPS-C 2015: Expressive Total, 4 = PEPS-C Receptive Total, 5 = TOWRE-2 SWE, 6 = TOWRE-2 PDE, 7 = TROG-2 Listening Comprehension, 8 = TROG-2 Reading Comprehension (Modified), 9 = YARC SWR, 10 = Passage Reading Comprehension (Raw Score), 11 = Passage Reading Comprehension (Standard Score). * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

Comprehension, YARC Passage Reading Comprehension (Raw Score), and YARC Passage Reading Comprehension (Standard Score). I will not discuss all of these outcomes below, but highlight outputs I think were worth talking about in more detail.

Beginning with the most significant deviation, the CTOPP-2 variable, the output of the statistic suggested a very significant non-normal distribution ($W = 0.914$, $p < 0.001$). But, taking this in conjunction with looking at the visual representations (see Appendix C.2.1), it is more clearly a negative skew, with a slight bimodal curve, as there is a dip in scores in the middle of the histogram. This is likely due to a difference in ability between children who understood the task more than others, as there was a stop condition for the CTOPP-2 Elision task which would occur if a child got three items in a row incorrect. As such, children who did not understand the instructions after repeated examples were likely to score low on the task.

The next variable that was highlighted was the TOWRE-2 SWE output ($W = 0.952$, $p < 0.05$). If I take the Shapiro-Wilk output as truth, then I have a non-normal distribution, but this would be a dangerous assumption to make without seeing the visualisation of the data. Looking at the histogram in Appendix C.6 I see a fairly normal distribution, which is affected by outliers are the lower end. These two outliers are discussed in Appendix C.4.1, in which I compared the means before and after removing the outliers. In terms of this variable, there was a small two-point change in mean with the removal of the outliers, which did not seem a significant change. As such, looking at the visualisation, the outlier seems to be two children who did not score as high on the task, and did not do anything wrong as such. Thus, for this variable I do not want to transform it, rather conduct the tests with and without the outliers, to see what the effect on the results is. I do not expect much change, but this will be discussed. As such, I chose not to transform this variable, believing that despite the indication of non-normality that the outliers were a valid aspect of the data and not a mistake, and were included in the analysis as they were.

Lastly, is the TROG-2 Reading Comprehension (Modified) variable ($W = 0.931$, $p < 0.01$). This measure was the most likely to be non-normal, because I was modifying an established measure to fit with our study, so it has not been widely tested with lots of samples to ensure it works perfectly. This can be seen with the wide range of scores and graphs in Figure C.2.8 and the boxplot. I tried some transformations, but the change was not enough to justify a full transformation. The distribution was affected by the scores on the tasks, I chose to keep it as it was.

Transformation Decisions

All of the above information led to decisions to be made regarding whether transforming the variables was appropriate and would happen. The full detail of these checks have been put into the appendices (see Appendix C.4.2, but as a summary transformations were conducted for the TROG-2 RC (Modified) variable, and the changes did improve the distribution, but I did not commit to this. I transformed the variable for the regression modelling through the use of Weighted-Least Squares modelling to reduce the heteroscedasticity of the residuals, which worked well for this. As such, I left the variable as it was and made adjustments through the method I used.

Chapter 4. Results

The final data set analysed was of the 51 children who had no known learning or neurological difficulties, out of the full cohort of 56 children. The structure of this chapter follows the order of tests discussed in Section 3.6.1, starting with the descriptive statistics for each task to provide a general overview of the average scores and spread of the data. Next, I performed multiple Pearson correlations between each task, and controlled for false discovery rate by using the Benjamini-Hochberg p-value adjustment method. Following this, I created multiple regression models to explore the research questions and hypotheses I proposed in Section 3.1. Finally, I created exploratory mediation models to further attempt to answer the research questions and hypotheses, with some caveats regarding causal inference in these models, discussed within Section 4.4.

I provided extra context to some sections, such as if transformations of data took place or data assumptions were critical for understanding. However, where appropriate some of this detail has been moved to the appendices, and throughout this chapter I reference where to find this extra information if the reader is interested to know more. Furthermore, the appendices also contain a hyperlink to the raw R scripts I used for the current research, for reference and transparency of approach.

4.1. Descriptive Statistics

4.1.1. Task Summaries

Figure 4.1 displays the descriptive statistics for all of the individual tasks used in the current project. Focusing on the prosody tasks, the output indicates that the Year 3 children scored similarly on the receptive (M = 82.7, SD = 9.7) and expressive (M = 83.4, SD = 10.4) tasks when totalled, when rounded both had a mean average of 83, which was unexpected due to the perception that the expressive tasks would be more difficult for the children compared to the receptive tasks. Looking closer at the individual tasks, there are no obvious tasks with extreme differences to indicate whether one was more difficult.

Moving focus to the decoding measures, and there is an expected difference between the TOWRE-2 SWE (M = 61.0, SD = 9.1) and PDE (M = 34.6, SD = 10.2) scores, in that the PDE task was a more difficult fluency task to complete due to the use of nonwords, and thus was harder for the participants. This is also reflected in the range of scores, with the minimum scored

	Mean	St. Dev.	Min	Max
PEPS-C 2015				
Expressive Tasks (/16)				
Affect Expressive	12.9	3.2	1	16
Boundary Expressive	11.3	3.0	3	16
Contrastive Stress Expressive	11.6	3.3	4	16
Imitation	12.5	2.3	7	16
Lexical Stress Expressive	9.5	2.6	5	16
Phrase Stress Expressive	11.7	2.2	6	15
Turn End Expressive	13.9	2.2	8	16
Receptive Tasks (/16)				
Affect Understanding	13.0	3.5	4	16
Boundary Understanding	12.5	2.8	7	16
Contrastive Stress Understanding	11.8	2.5	6	16
Discrimination	11.4	2.8	5	16
Lexical Stress Understanding	10.3	2.3	6	14
Phrase Stress Understanding	9.5	2.6	4	16
Turn End Understanding	14.2	2.5	6	16
Totalled Tasks				
Receptive Tasks (/112)	82.7	9.7	63	104
Expressive Tasks (/112)	83.4	10.4	58	103
Totalled Tasks (/224)	166.0	16.8	127	200
Decoding Measures				
TOWRE-2 SWE (/108)	61.0	9.1	39	78
TOWRE-2 PDE (/66)	34.6	10.2	10	58
YARC SWR (/60)	40.8	7.3	25	52
Listening Comprehension				
TROG-2 LC (/42)	33.5	3.1	27	39
Reading Comprehension				
YARC Passage Reading Comprehension (/16)	9.7	2.9	5	15
YARC Passage Reading Standard Scores	105.3	8.9	91	124
TROG-2 RC (Modified) (/42)	30.7	3.9	19	39
Control Tasks				
CTOPP-2 Elision (/34)	25.2	5.4	13	33
Backwards Digit Span Raw Score	14.4	4.0	7	26

Table 4.1 Descriptive statistics for the raw scores of each task in the project (N=51).

on the SWE task 39, and the PDE task this was 10, though the max score for each was different so further interpretations should be taken with caution. The other decoding measure I used was the YARC SWR (*M* = 40.8, *SD* = 7.3), a more traditional word recognition task which did not have the time pressure of the previous tasks.

The TROG-2 measures were matched to use the same stimuli, except the nouns were swapped in each sentence. It is interesting to see that the TROG-2 LC task had a slightly higher mean (*M* = 33.5, *SD* = 3.1) compared to the TROG-2 RC (*M* = 30.7, *SD* = 3.9). This difference is not too extreme, but looking at the ranges of scores there is a wider gulf, with the TROG-2 LC range from 27-39, and the TROG-2 RC range from 19-39.

Moving to the YARC Reading Comprehension scores, there are a range of scores between 5 - 15 out of 16, showing a range of abilities when answering these comprehension questions. Looking at the standard scores next the mean (*M* = 105.3, *SD* = 8.9) suggests the children scored around average for their age group.

Finally, the control task data. For the CTOPP-2 Elision task there was a range of scores from 13 - 33, not quite a ceiling effect but close. On average the children scored 25/34 on the elision task. For the backwards digit span task, there was a wide range of scores on this task from 7 - 26 number strings correctly identified, with an average of 14 correctly remembered.

4.2. Correlations

The descriptive statistics gave us a general overview of the outline of our data, but now we are now moving into investigating the relationships between the different tasks. It is important to note at this point that going forward I am using the totalled versions of the PEPS-C tasks. I am most interested in the differences between the receptive and expressive tasks (if there are any) in this project, so it makes sense to take these totalled amounts forward for analysis. All the results discussed below take their values from Figure 4.1. The correlellogram highlights only the correlatory relationships that were significant with colour, so any cells that are not filled with colour are non-significant. I also chose the alpha value of 0.01 for significance, as I wanted to try and reduce the number of false positive relationships in these relationships.

The coefficient relationships were all positive in direction (except for between Age Months and Digit Span, but this was small at -0.02) suggesting that all the tasks I chose to measure had a positive linear relationship of varying strengths. Furthermore, Age Months was not significantly correlated with any variable, so I did not use this variable for the forthcoming regression modelling or mediation model.

The first hypothesis proposed for the project was that the measures of prosody would have a significant relationship with the measures of reading comprehension. Looking at the correlation coefficients we see that the PEPS-C Expressive total correlated moderately with the YARC Reading Comprehension task (r = .45) and the TROG-2 Modified Reading Comprehension task (r = .57). The Receptive totalled tasks correlate moderately with the YARC Reading

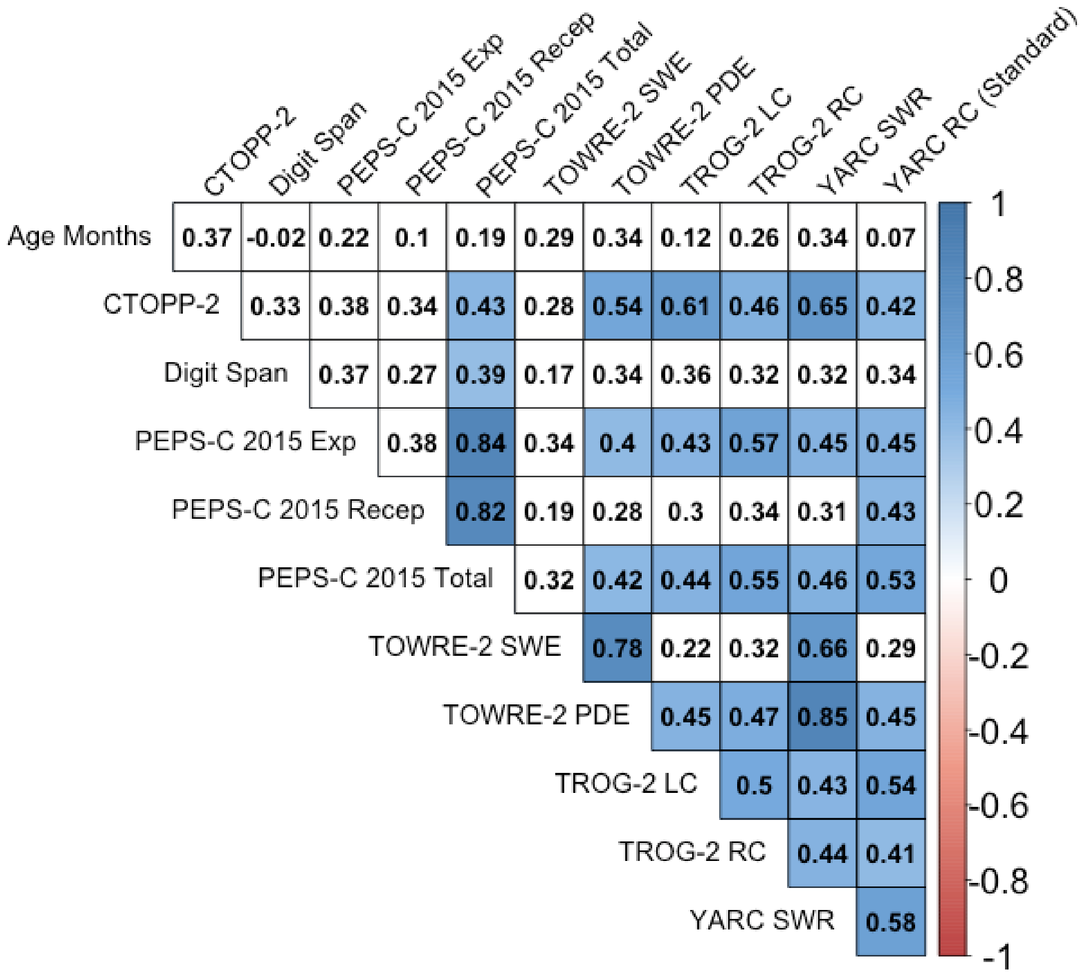

Figure 4.1 Pearson Correlation Plot: Each task has been correlated with the other task in this plot. The colour of the individual squares indicates the strength of the directional relationship between the two variables. If any square has no colour, it means the p-value was not less than 0.01. Created using the *corrplot* package created by Wei and Simko (2017).

Comprehension task (r = .43), but the correlation coefficient was lower for the TROG-2 Reading Comprehension task (r = .34). I also had the variable for the PEPS-C 2015 tasks totalled, and this showed a moderate correlation with the YARC RC task (r = .53) and a moderate correlation with the TROG-2 Reading Comprehension task (r = .55).

The correlation between the TOWRE-2 SWE and the other measures was insignificant and in the low range, except for strong correlations with the other measures of decoding (TOWRE-2 PDE, r = .78; YARC Short Word reading, r = .66). Comparatively, the TOWRE-2 PDE task had a moderate to strong correlation with many of the tasks, including the PEPS-C Expressive tasks (r = .40) and the totalled variable (r = .42), though not the Receptive tasks alone (r = .28). Furthermore there were moderate correlations with all of the reading comprehension tasks (YARC RC Standard Score, r = .45; TROG-2 RC, r = .47), and the listening comprehension task (r = .50). The final task in our decoding set was the YARC Short Word Reading task, taken from the YARC Passage Reading set of tasks. Unsurprisingly, this task was correlated with the YARC Reading Comprehension task (Raw Score, r = .64; Standard Score, r = .58), and moderately correlated with the other measure of reading comprehension (TROG-2 RC, r = .44). There was also a very high correlation between performance on the YARC SWR task and the TOWRE-2 PDE with a coefficient of r = .85.

The other major task in the set was the TROG-2 Listening Comprehension, the last component of the Simple View of Reading. Looking at the relationship with reading comprehension we see that LC had a moderate relationship with all the tasks, which was at least expected with the TROG-2 Reading Comprehension task due to both using the same edited stimuli (YARC RC Standard Score, r = .54; TROG-2 RC, r = .50).

Finally, a short summary of the relationship between the prosody tasks and the measures of the SVR. Taking the Expressive tasks first, we see a moderate relationship with each decoding task, with only the TOWRE-2 SWE showing non-significance at the 0.01 alpha level (TOWRE-2 SWE, r = .34; TOWRE-2 PDE, r = .40; YARC SWR, r =.45). Regarding the listening comprehension part of the SVR framework we see as well a moderately significant relationship between scores on the Expressive tasks and the TROG-2 Listening Comprehension task (TROG-2 LC, r = .43). Comparatively, the Receptive tasks showed lower coefficient relations with the SVR tasks, and none at the 0.01 alpha level I set (TOWRE-2 SWE, r = .19; TOWRE-2 PDE, r = .28; TROG-2 LC, r = .30; YARC SWR, r = .31).

4.3. Regression Analysis

This section presents the output of the many multiple regression models conducted to investigate prosody's relationship with various measures from the project. The plan for the regression models was detailed in Section 3.6.1, but as a summary there were eight models of multiple regression created to investigate the various research questions and hypotheses I had, beginning with looking at the direct relationship of the receptive and expressive prosody totalled scores with the aspects of speech and reading discussed in these questions (reading comprehension, decoding and linguistic comprehension). Following this there are two multiple regression models with all the tasks as predictors regressed onto the individual reading comprehension tasks, which was to see if prosody's relationship with reading comprehension was softened by this co-occurrence with the other measures. Each model is analysed in turn, providing a summary of the model output, which will be discussed and critiqued in the following chapter.

For the models where I used all the tasks as predictors, I standardised the predictors by transforming them into z-scores. This procedure is described in Subsection 4.3.2, alongside specific issues regarding multicollinearity and how this led to the creation of a decoding composite variable for use in the models.

4.3.1. Multiple Regression Models: Prosody Measures as Predictors

YARC Passage Reading

The first multiple regression model was to look at the contribution of the composite prosody variables to the outcome variable of the YARC Passage Reading (Standard Score). The model met the assumptions for a multiple regression, so no transformation occurred.

	Dependent variable:
	YARC Passage Reading (Standard Score)
PEPS-C 2015: Expressive	0.284^{**}
	(0.114)
PEPS-C 2015: Receptive	0.276^{**}
	(0.122)
Constant	58.828^{***}
	(10.937)
Observations	51
R^2	0.277
Adjusted R^2	0.247
Residual Std. Error	7.751 (df = 48)
F Statistic	9.179^{***} (df = 2; 48)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 4.2 Multiple Regression Model 1: Model of the relationship between the predictors of the PEPS-C 2015 test and the outcome YARC Passage Reading (Standard Scores).

The final regression model equation was as so: $\hat{Y}$ = 58.828 + 0.284 (PEPS-C 2015: Expressive) + 0.276 (PEPS-C 2015: Receptive). The regression model was significant, indicating that the PEPS-C 2015 Expressive and Receptive task combined predicted the standard scores for the YARC Reading Comprehension task (F(2, 48) = 9.179, p < 0.001, $R^2 = 0.28$). Both predictors were individual predictors of standard scores on their own terms (Expressive: $\beta = 0.28$, t(48) = 2.262, p = 0.0160, $pr^2 = .11$; Receptive: $\beta = 0.28$, t(48) = 2.262, p = 0.0283, $pr^2 = .10$).

TROG-2 Modified Reading Comprehension Measure

The same procedure was followed as 4.3.1 to look at the relationship between prosody and this measure of reading comprehension (($\hat{Y}$ = 10.159 + 0.183 (PEPS-C 2015: Expressive) + 0.064 (PEPS-C 2015: Receptive)). This is seen in Table 4.3.

When checking the assumptions for this model, I found an issue with heteroscedasticity with the residuals. I found the residuals showed a 'fan-shape' when visualising them, suggesting this was an issue. To overcome this issue I used a Weighted Least Squares (WLS) model for this relationship. WLS regression is used in cases where heteroscedasticity is an issue, it helps to correct non-constant variance by weighting each of the observations by the reciprocal of its estimated variances. To calculate the weights, I coded a formula in R wherein I estimated the weight of the model, which was estimated from residual plots. We are not transforming the model, rather treating each observation as more or less informative, giving appropriate 'weight' to the informativeness of the values.

All other assumptions were met without issue, and in re-checking the assumptions after re-fitting the model I found I met the assumption of homoscedasticity, with minimal change to output of the model.

	Dependent variable:
	TROG-2 Reading Comprehension
PEPS-C 2015: Expressive	0.183^{***}
	(0.041)
PEPS-C 2015: Receptive	0.064
	(0.040)
Constant	10.159^{**}
	(4.343)
Observations	51
R^2	0.371
Adjusted R^2	0.345
Residual Std. Error	1.366 (df = 48)
F Statistic	14.145^{***} (df = 2; 48)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 4.3 Multiple Regression Model 2: Model of the relationship between the predictors of the PEPS-2015 test and the outcome TROG-2 Reading Comprehension (modified). This model was conducted using the WLS regression method to help control for the issues with heteroscedasticity.

The model shown in Table 4.3 suggests an overall significant relationship between the predictors of prosody and the outcome of TROG-2 Reading Comprehension (F(2,48) = 14.145, p < 0.01, R^2 = 0.37). Looking next at the individual predictors, we see the PEPS-C 2015: Expressive tasks significantly predicted scores on the TROG-2 Reading Comprehension task (β = 0.183, , t(48) = 4.448, p < 0.01, pr^2 = 0.0030). Comparatively, the PEPS-C 2015: Receptive predictor was not significantly predictive of TROG-2 Reading Comprehension scores (β = 0.064, t(48) = 1.597, p > 0.1, pr^2 = 0.0077).

Prosody and the Simple View of Reading Measures

Before looking the multiple regression model, I modelled the relationship between the measures of prosody and the components of the SVR. I hypothesised that the measures of prosody would be related to the measure of linguistic comprehension I chose, as prosodic understanding is a key component of understanding speech and garnering meaning from the speech stream. Comparatively, the relationship between prosody and decoding does exist (prosody can help with stress placement in words for example), but is not expected to be as key to this skill set.

Linguistic Comprehension

I created a multiple regression model with the measures of prosody as predictors, and the TROG-2 Listening Comprehension raw score output as the outcome ($\hat{Y}$ = 20.201 + 0.111 (PEPS-C 2015: Expressive) + 0.049 (PEPS-C 2015: Receptive)). The model output suggested that the measures of prosody were significant predictors of listening comprehension ability ($F(2,48) = 6.296$, $p < 0.01$, $R^2 = 0.208$). Looking at the individual predictors, we see that the PEPS-C Expressive score was a highly significant contributor to the listening comprehension outcome ($\beta = 0.111$, $t(48) = 2.685$, $p < 0.01$, $p^2 = 0.1305$). The PEPS-C Receptive score was not a significant predictor of listening comprehension outcome ($\beta = 0.049$, $t(48) = 1.117$, $p > 0.1$, $p^2 = 0.0253$).

	Dependent variable:
	TROG-2 LC
PEPS-C Expressive	0.111^{***}
	(0.041)
PEPS-C Receptive	0.049
	(0.044)
Constant	20.201^{***}
	(3.958)
Observations	51
R^2	0.208
Adjusted R^2	0.175
Residual Std. Error	2.805 (df = 48)
F Statistic	6.296^{***} (df = 2; 48)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 4.4 Multiple Regression 3: Model with the measures of prosody as predictors, and the measure TROG-2 listening comprehension measure as the outcome variable.

Decoding Measures

The next three multiple regression models investigate the relationship of the PEPS-C variables with the three measures of decoding used in the current project. The results of these models are displayed in Table 4.5. Each model was screened for outliers, with only the TOWRE-2 SWE model highlighting possible errors. However, I chose to include the two variables I identified as potential outliers, because they were valid values scored by the task and did not appear to affect the output of the results significantly. Similarly, all assumptions for conducting a regression were met.

The multiple regression model to predict the TOWRE-2 SWE raw score used the PEPS-C expressive and receptive measures are predictors ($\hat{Y}$ = 32.485 + 0.272 (PEPS-C 2015: Expressive) + 0.070 (PEPS-C 2015: Receptive)). The overall model was significant for predicting the score on the TOWRE-2 SWE task via the measures of expressive and receptive prosody (F(2, 48) = 3.26, $p < 0.05$, $R^2 = 0.12$). When looking at the predictors however, it appears that only the Expressive Prosody predictor was significantly predictive of TOWRE-2 SWE raw score outcome ($\beta = 0.27230$, t(48) = 2.629, p = 0.0394, $pr^2 = 0.085$). Comparatively, the Receptive Prosody tasks did not significantly predict TOWRE-2 SWE raw score outcome ($\beta = 0.06989$, t(48) = 2.118, p = 0.6143, $pr^2 = 0.0053$).

The next model is the modelling of the prosody measures predicting the outcome of the TOWRE-2 PDE task ($\hat{Y}$ = -6.617 + 0.339 (PEPS-C 2015: Expressive) + 0.157 (PEPS-C 2015: Receptive)). The model overall was significant indicating that the predictors of PEPS-C 2015 Expressive and Receptive were significantly predictive of values for TOWRE-2 PDE (F(2, 48) = 5.374, $p < 0.01$, $R^2 = 0.183$). When looking further at the predictors themselves, we see that like the TOWRE-2 SWE model that only the PEPS-C Expressive total variable was significantly predictive of non-word reading ability ($\beta = 0.3391$, t(48) = 2.458, p = 0.0176, $pr^2 = 0.1128$). The PEPS-C Receptive total variable on the other hand was not significantly predictive of non-word reading ability ($\beta = 0.1572$, t(48) = 1.063, p = 0.2929, $pr^2 = 0.0230$).

The final model for this set is the modelling of the YARC SWR outcome ($\hat{Y}$ = 8.345 + 0.268 (PEPS-C 2015: Expressive) + 0.122 (PEPS-C 2015: Receptive)). The model was significant overall, indicating that the presence of PEPS-C Expressive and Receptive contributed towards the YARC SWR score in some capacity (F(2,48) = 6.913, $p < 0.01$, $R^2 = 0.224$). Looking closer at the individual predictors, we see that the PEPS-C Expressive task scores were significantly predictive of the YARC SWR scores ($\beta = 0.272$, t(48) = 2.800, $p < 0.01$, $pr^2 = 0.1404$). PEPS-C Receptive task scores were not significantly predictive of the YARC SWR scores according to our model ($\beta = 0.122$, t(48) = 1.190, $p > 0.1$, $pr^2 = 0.0286$).

4.3.2. Multiple Regression Models: All Variables as Predictors

The next stage of regression modelling is to model the relationship with the reading comprehension variables, including all the variables I measured in the project. This includes the measures representing the framework of the SVR (TOWRE-2 SWE, TOWRE-2 PDE, TROG-2 LC, YARC SWR), but also our control measures (CTOPP-2 Elision, Backwards Digit Span), to see if the measures of prosody still provide a significant contribution to our models when the other variables are present. I discuss below some of the preliminary steps I took to create these

	Dependent variable:		
	TOWRE-2 SWE	TOWRE-2 PDE	YARC SWR
	(1)	(2)	(3)
PEPS-C Expressive	0.272^{**}	0.339^{**}	0.268^{***}
	(0.129)	(0.138)	(0.096)
PEPS-C Receptive	0.070	0.157	0.122
	(0.138)	(0.148)	(0.103)
Constant	32.485^{**}	-6.617	8.345
	(12.357)	(13.259)	(9.210)
Observations	51	51	51
R^2	0.120	0.183	0.224
Adjusted R^2	0.083	0.149	0.191
Residual Std. Error (df = 48)	8.758	9.397	6.527
F Statistic (df = 2; 48)	3.260^{**}	5.374^{***}	6.913^{***}
Note:			$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 4.5 Multiple Regression Model(s) 4: Table of the three multiple regression models with the measures of prosody as predictors, and the three measures of decoding ability as outcomes.

models, including standardising the predictors to make them more comparable on the same scale, and reflecting on issues with multicollinearity and the related decision I made to create a composite variable to deal with this.

Standardising the Predictors

Before conducting the multiple regression analysis I decided to standardise the predictor variables into z-scores. I believed that because the scales of measurement varied between the tasks so much that it would be better to compare between different tasks using z-scores compared to the raw data by transforming these scores. To do this I first centred each variable by subtracting the mean from each child's individual score. Following this, I then took the centred scores and divided them by the standard deviation of the centred score, to get our final centred z-scores for each variable. This means our transformed values continue to have the same relationship to one another, but the transformed variable has a mean of zero and a standard deviation of one. It is important to note that the underlying model is the same as it would be using the unstandardised predictors, but the way we interpret the output changes, such that we discuss how much the outcome changes by a one standard deviation change in the predictor.

Multicollinearity in the Modelling

One of the issues that arises with the inclusion of multiple predictors in regression models is multicollinearity. In our modelling I used the Variance Inflation Factor (VIF) to assess the severity of multicollineairty in our models, assessing whether I included too many variables in our model. This became an issue when I input all of our decoding variables (TOWRE-2 SWR, TOWRE-2 PDE, and YARC SWR) into the model, and the VIF values were noticeably higher, suggesting I was over-inflating decoding's contribution to model, and this concept could be represented by one or two of the variables instead of all of them in the model. I decided to create a composite measure of decoding ability to represent each of these variables as one, because of the theoretical underpinning I followed regarding the SVR it was suitable to combine these measures of decoding into one. Furthermore, because of the small sample and number of predictors in this model, the composite variable helped to improve the power of the model with the lessening of multiple comparisons.

To create the composite variable I first centred each variable, then I standardised each individual score. Following this I then used R to calculate the average of each participant's decoding scores, essentially the average z-score between the TOWRE-2 SWE, TOWRE-2 PDE and YARC SWR variables. The output of these three steps gave us the final Decoding Composite variable. An important note to remember is our interpretation of this construct in the model, it is a composite of the decoding measures and needs to be treating as such, it is a representation and not a pure measure. That said, as above, this fits with the SVR framework as a representation of decoding, so I felt it was suitable.

YARC RC Standard Scores

The first model is the standardised predictors regressed against the YARC Passage Reading Comprehension Standard Score outcome ($\hat{Y}$ = 105.31 + -0.51 (CTOPP-2 (Z-Score)) + 2.37 (Decoding Composite) + 0.52 (Backwards Digit Span (Z-Score)) + 0.95 (PEPS-C 2015: Expressive (Z-Score)) + 1.93 (PEPS-C 2015: Receptive (Z-Score)) + 3.13 (TROG-2 LC (Z-Score)). Each variable was entered simultaneously into the model. When looking at the outliers I used the Mahalanobis scores, measures of leverage and the Cook's Distance measure. I then collated how many participant's scores violated the cutoff points for these tests. I found that two participant's outputs had leverage slightly above our cutoff point of 0.117 and above our Cook's Distance cutoff of 0.083. I decided to keep these outlier scores in the final model, and this issue may have arisen due to a small sample size. This is discussed more in Section 5.3.1 in the following chapter.

When conducting assumptions checks I found issues with multicollinearity discussed above. I then conducted the standardising procedure for the Decoding Composite measure, and re-ran the Variance Inflation Factor measure, and found that the multicollinearity issue was settled. All other assumptions to conduct the regression model were met successfully.

The results of the regression model can be seen in Table 4.6. The overall model was highly significant in predicting the outcome variable of YARC Passage Reading Standard Scores ($F(6,44) = 10.04$, $p < 0.01$, $R^2 = 0.58$). Looking closer at the individual predictors I see that the only significant predictor of YARC Passage Reading scores was the TROG-2 LC variable ($\beta = 3.13$, $t(44) = 2.36$, $p < 0.05$, $pr^2 = 0.11$). Two other variables were also notable, with larger coefficients than the other variables, but not at the 0.05 alpha value commonly used to identify significance in these tasks. The first was the Decoding Composite variable ($\beta = 2.37$, $t(44) = 1.74$, $p < 0.1$, $pr^2 = 0.06$), and secondly the PEPS-C 2015 Receptive tasks ($\beta = 1.93$, $t(44) = 1.72$, $p < 0.1$, $pr^2 = 0.06$).

	Dependent variable:
	YARC Passage Reading (Standard Score)
CTOPP-2 (Z-Scores)	-0.51
	$(-3.25, 2.22)$
Decoding Composite	2.37^{*}
	$(-0.30, 5.04)$
Backwards Digit Span (Z-Score)	0.52
	$(-1.69, 2.74)$
PEPS-C 2015: Expressive (Z-Score)	0.95
	$(-1.47, 3.37)$
PEPS-C 2015: Receptive (Z-Score)	1.93^{*}
	$(-0.27, 4.13)$
TROG-2 LC (Z-Score)	3.13^{**}
	$(0.53, 5.74)$
Constant	105.31^{***}
	$(103.36, 107.26)$
Observations	51
R^2	0.44
Adjusted R^2	0.37
Residual Std. Error	7.10 (df = 44)
F Statistic	5.83^{***} (df = 6; 44)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 4.6 Multiple Regression Model 5: Output for the multiple regression model predicting the outcome scores for YARC Passage Reading (Standard Score). Shown above are the coefficients for each variable in our study, which have been centered and standardised for legibility in comparing the effects of each task, i.e. for one standard deviation change in TROG-2 LC, YARC Passage Reading Standard scores are predicted to increase by 3.13 points. Furthermore, I included 95% confidence intervals below each coefficient, as an extra measure to check contribution.

TROG-2 Modified: Reading Comprehension

The last multiple regression model I created was the simultaneous multiple regression model, including all of our test and control variables, to identify the contribution of prosody to the TROG-2 Modified RC variable once the other tasks were included ($\hat{Y}$ = 30.65 + 0.36 (CTOPP-2 (Z-Score)) + 1.06 (Decoding Composite) + -0.30 (Backwards Digit Span (Z-Score)) + 1.17 (PEPS-C 2015: Expressive (Z-Score)) + 0.52 (PEPS-C 2015: Receptive (Z-Score)) + 1.21 (TROG-2 LC (Z-Score)).

As identified when TROG-2 RC was an outcome variable prior, the assumption of homoscedasticity was not met when testing the model met the regression assumptions. Furthermore, when looking at the standardised residuals of the model I found the normality assumption was not met either. To overcome this I again used the Weighted Least Squares general linear model to account for the heteroscedasticity present in the model. Below I again present the original model under the OLS heading, and the final model under the WLS heading after applying the weights to the model. This led to the data meeting the assumptions of homoscedasticity and normality.

Furthermore, I also used the above composite variable for decoding once more, as again the VIF formula and correlatory output identified high instances of correlation between the measures of decoding. Furthermore, all outputs were centred and standardised for easier interpretation between the tasks and the size of the effect.

From the output of Table 4.7 we see the two models side-by-side, these correspond to the original Ordinary Least Squares model on the left, and the Weighted Least Squares model on the right. I provided both in this case as a comparison, showing the differences that came with the adding of the weighting. I choose to summarise the WLS results in this case as this modelling helped to correct for heteroscedasticty and normal distribution. The WLS model surmises that the predictors included in this model were able to significantly predict the score on the TROG-2 Reading Comprehension (Modified) task (F(6,44) = 10.39, p < 0.01, R^2 = .59). Next, looking at the predictors in the model we can see that there was a single predictive variable that was significant at the 95% level, and that was the PEPS-C 2015: Expressive tasks raw score (β = 1.16, t(44) = 2.556, p < 0.05, pr^2 = 0.12). This was similar to the original OLS model too, though with a slight change to the unstandardised beta value and the confidence intervals. There were no other highly significant predictors identified by this model, the closest were the Decoding Composite measure (β = 1.04, t(44) = 1.708, p < 0.01, pr^2 = 0.022) and the TROG-2 Listening Comprehension measures (β = 1.06, t(44) = 1.965, p < 0.01, pr^2 = 0.03).

	Dependent variable:	
	TROG-2 Reading Comprehension (Modified)	
	OLS	WLS
	(1)	(2)
CTOPP-2 (Z-Score)	0.37	0.36
	(-0.83,1.57)	(-0.78,1.50)
Decoding Composite	0.58	1.06**
	(-0.59,1.75)	(0.04,2.08)
Backwards Digit Span (Z-Score)	0.12	−0.30
	(-0.85,1.09)	(-1.05,0.45)
PEPS-C 2015: Expressive (Z-Score)	1.33**	1.17**
	(0.27,2.39)	(0.30,2.04)
PEPS-C 2015: Receptive (Z-Score)	0.24	0.52
	(-0.72,1.21)	(-0.22,1.26)
TROG-2 LC (Z-Score)	0.84	1.21**
	(-0.31,1.98)	(0.20,2.23)
Constant	30.71***	30.65***
	(29.85,31.56)	(29.81,31.49)
Observations	51	51
R^2	0.44	0.60
Adjusted R^2	0.36	0.55
Residual Std. Error (df = 44)	3.11	1.39
F Statistic (df = 6; 44)	5.71***	11.15***
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 4.7 Multiple Regression Model 6: Output for the multiple regression model of predicting the score for the TROG-2 Reading Comprehension (Modified) raw scores. Shown above are the original Ordinary Least Squares model, and the Weighted Least Squares model which accounted for the issue of heteroscedasticity by using an estimate of the standard deviation of the original models values and met the assumptions for the regression. Please note that the coefficients in this table are based upon centred and standardised predictors, meaning the model is interpreted in how a change one standard deviation affects the comprehension score, e.g. a 1SD change in PEPS-C 2015 Expressive task score predicted a 1.16 increase in TROG-2 Reading Comprehension score.

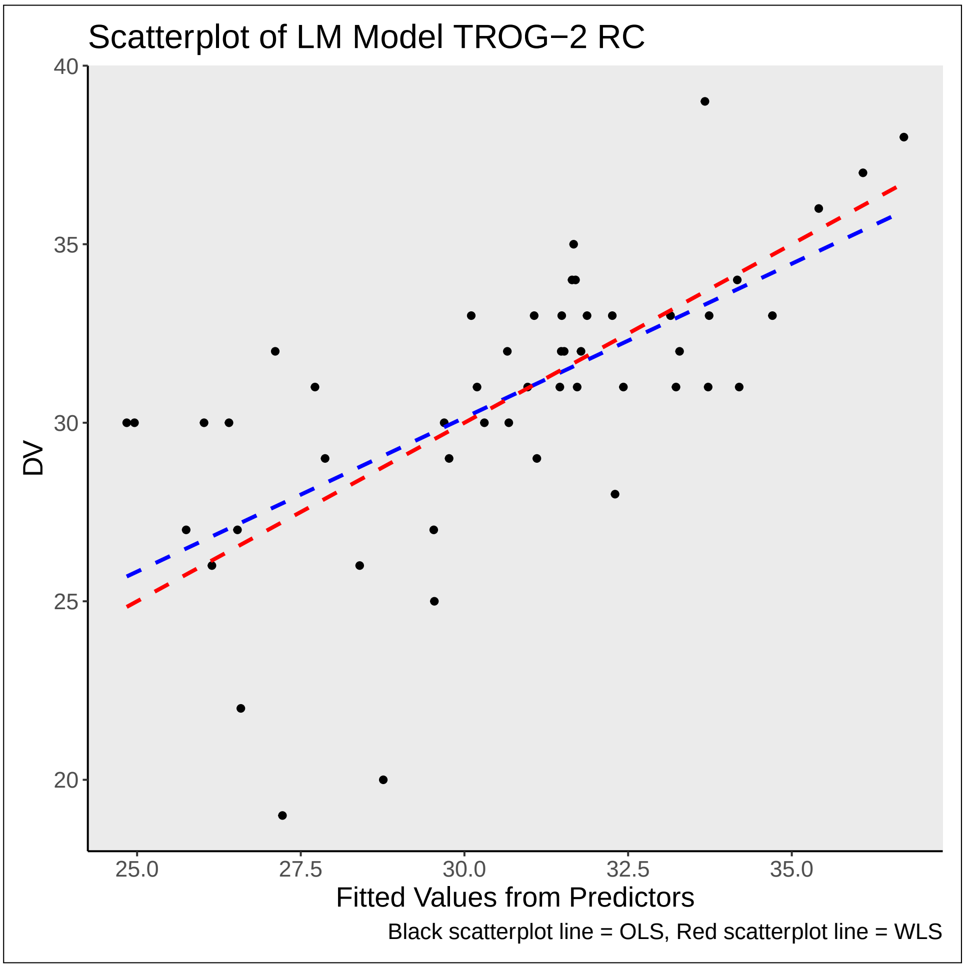

Figure 4.2 Scatterplot of fitted vs. actual values from the TROG-2 Reading Comprehension multiple regression model. The WLS Regression line has been weighted to fit more of the data points, and resolve the issues with heteroscedasticity found prior.

4.4. Mediation Analysis

The mediation model technique was chosen to assess the possible independent contribution of prosody ability with reading comprehension, controlling for the components of the SVR to see whether partial or full mediation occurred.

The original method for investigating mediation models was the Baron and Kenny (1986) four step method. In short, the researcher goes through four steps to determine whether a mediator effect has occurred, which are as so:

- *Step 1:* Estimate the relationship between the predictor (X) and the outcome (Y). Essentially, conduct a linear regression. this is known as Path c, and they must be significantly different from zero, showing an effect exists between the two variables.
- *Step 2:* Estimate the relationship between the predictor (X) on the mediator (M). Again, this is a basic linear regression to look at the relationship between the two. This is path a, and it also must be significantly different from 0, showing a relation exists between the two.

- *Step 3:* Estimate the relationship between the mediator (M) on the outcome (Y), whilst controlling for the predictor (X). This is a multiple regression model. This is known as path b, and again is must be significantly different from 0, the mediator and outcome must be related.

- *Step 4:* Finally, estimate the relationship between the outcome (Y) on the predictor (X), whilst controlling for the mediator (M). This should be non-significant, and near 0. This is known as reversed path c'.

A simple visualisation of these pathways can be seen in Figure 4.3, which labels each pathway as above.

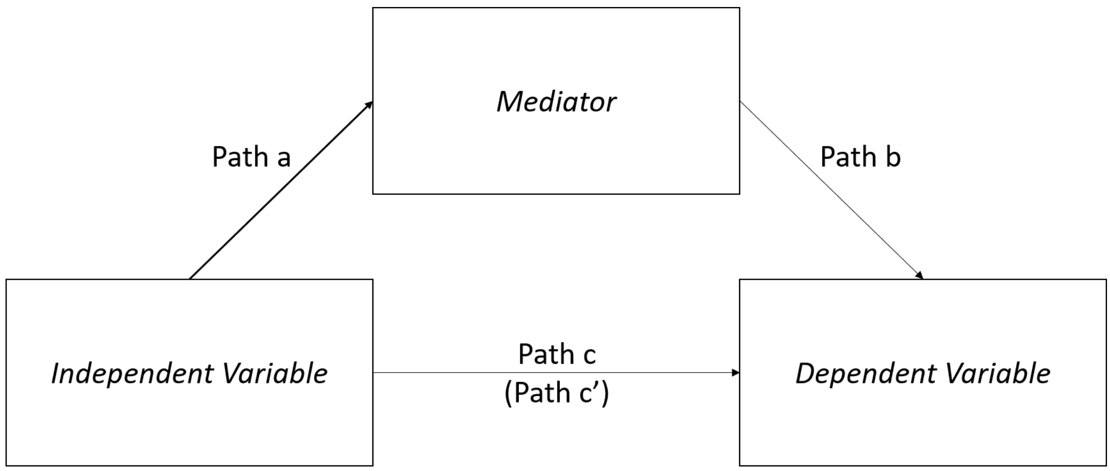

Figure 4.3 Simple Mediation Model Visualisation: Diagram for the simple mediation design, with the pathways from the Baron and Kenny (1986) four step method. Each pathway is a separate linear regression model, which are put together to assess mediation.

Once these regressions have been completed and analysed, the researcher uses the Sobel Test, a specialised t-test, to investigate these pathways and investigate mediation. However, the Sobel Test has fallen out of favour in recent years, related to issues with power and its assumption that the indirect effect path (path ab) is normally distributed. Thus, current users of mediation have move towards using a bootstrapping method instead, most notably discussed by Preacher and Hayes (2004), who created a macro called PROCESS in SPSS that became widely used for this purpose. We can now do this same procedure within the R statistical software. We compute two pathways for the mediation model:

- *Step 1:* Create a linear regression between the predictor (X) and the mediator (M). This is path "a".

- *Step 2:* Create a multiple regression model between the predictor (X) and the outcome (Y), controlled by the mediator (M). This is path "b"

Taking these paths together, we get an estimate of the indirect effect (path "ab"), and are able to sample this using bootstrapping procedures (this tends to be 1000 times, but this can be edited within the R code). We can run the mediation model with and without the bootstrapping, to verify the results with increased power. This is the basic model I used to help answer our research questions, though I will still be referring to these other paths in the output, the important pathway is the indirect one to assess whether any mediation has occurred.

An important note to make at this point, and which will be discussed in our discussion section, is that mediation analysis is used often as a causal explanation of the relationship between two variables. As such, there would be rightful queries about this project using mediation analysis to assess causality at one time point. However, I have decided to use mediation analysis as an exploratory method to look at this data, providing suggestions for further work with multiple age groups or over time to implement these models. This was inspired by Hayes (2018) who discussed that even when the data are not unequivocally suitable for causal claims, it is on the researcher to make clear these caveats, but also for the researcher to "interpret and place meaning on the mathematical procedures used, not the procedures themselves" (p.81). This will be discussed further in the discussion chapter, but I felt using a mediation model was an interesting way to estimate and examine the relationship between prosody, reading comprehension and the SVR components.

For the current data I wanted to use mediation modelling to help answer research question two, assessing whether prosody directly relates to reading comprehension after controlling for decoding and linguistic comprehension. To do this I used R once more and the *mediation* package (Tingley et al., 2014), a package which takes the set of regressions for a mediation and fits them for the user. The *mediation* package outputs the results of the mediation model with a set of results with acronyms that stand for certain aspects, and which mean certain things in the model. Below is a list of these acronyms and titles with an explanation:

- ACME = Average Causal Mediation Effects: this is the indirect effect of the predictor on the outcome through the mediator. This is the "ab" path which was discussed above, and is the pathway computed by Preacher and Hayes (2004) in their PROCESS function, and similarly in this model I bootstrapped the mediation model with 1000 replications to check the pathway was still significant.

- ADE = Average Direct Effects: this is the direct effect of the predictor on the outcome. The coefficient recorded here is the output from the multiple regression computed when the mediation variable was included, so it is the direct effect when controlled for by the mediator.

- Total Effect: this is as the name suggests, the combined effect of the indirect and direct pathways, computed by adding them together.

- Proportion Mediated: the final aspect of the mediation output is the proportion of the effect of the predictor on the outcome that goes through the mediator according to the functions output. This is calculated by dividing the ACME through the total effect. This is an interesting piece of extra information provided by the function, but it should not be the focus of the analysis, and as such I only mention it in passing.

I created various models, one for the PEPS-C 2015 Expressive scores and one for the PEPS-C 2015 Receptive scores. Following on from the previous subsections, I wanted to explore the difference in contribution between the two types of prosody ability identified in the PEPS-C 2015 tasks. As such, I have created models for each of these scores. Furthermore, I have created a composite decoding skill containing the measures of the TOWRE-2 and YARC word reading subtests as seen in the regression section. Furthermore, I ran the models with bootstrapping in an attempt to increase the power of the output. Each model was programmed to run 1000 replications of random samples from the data, and output 95% confidence intervals for each output. These confidence intervals are shown in brackets for each measure in Table 4.8.

It is worth noting here that I did not create a multiple mediation model for this project, though the logic is there to conduct such a model to answer our research question I did not meet the power requirements to run such a model due to our sample size. This issue will be discussed further in our Discussion chapter, and come as part of our suggestions for further study into this area. For the moment, I focus upon the mediation models created below.

	Mediation Models: YARC Passage Reading			
	(1)	(2)	(3)	(4)
ACME	1.67^{***}	1.22^{***}	1.35^{**}	0.99^{**}
	(0.56, 3.17)	(0.35, 2.29)	(0.23, 2.56)	(0.07, 2.12)
ADE	2.32^{**}	2.60^{**}	2.64^{**}	2.83^{**}
	(0.18, 4.70)	(0.26, 4.82)	(0.39, 5.39)	(0.55, 5.12)
Total Effect	3.99^{***}	3.82^{***}	3.99^{***}	3.82^{***}
	(1.70, 6.19)	(1.27, 5.87)	(1.70, 6.19)	(1.27, 5.87)
Proportion Mediated	0.42^{***}	0.32^{***}	0.34^{**}	0.26^{**}
	(0.14, 0.92)	(0.09, 0.85)	(0.06, 0.82)	(0.02, 0.71)
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$			

Table 4.8 Table of the mediation pathway outputs for each model, with the independent variable as either PEPS-C 2015 expressive prosody or receptive prosody, mediator variable was either the decoding composite or TROG-2 listening comprehension task, and finally the dependent variable was the YARC Passage Reading (Standard Scores). See the individual model diagrams for the visualisation of the below results. (1) = Mediation Model 1 (Figure 4.4), (2) = Mediation Model 2 (Figure 4.5), (3) = Mediation Model 3 (Figure 4.6), (4) = Mediation Model 4 (Figure 4.7). ACME = Average Causal Mediation Effects, ADE = Average Direct Effects.

4.4.1. Mediation Model One: Expressive Prosody and YARC Passage Reading Mediated by Listening Comprehension

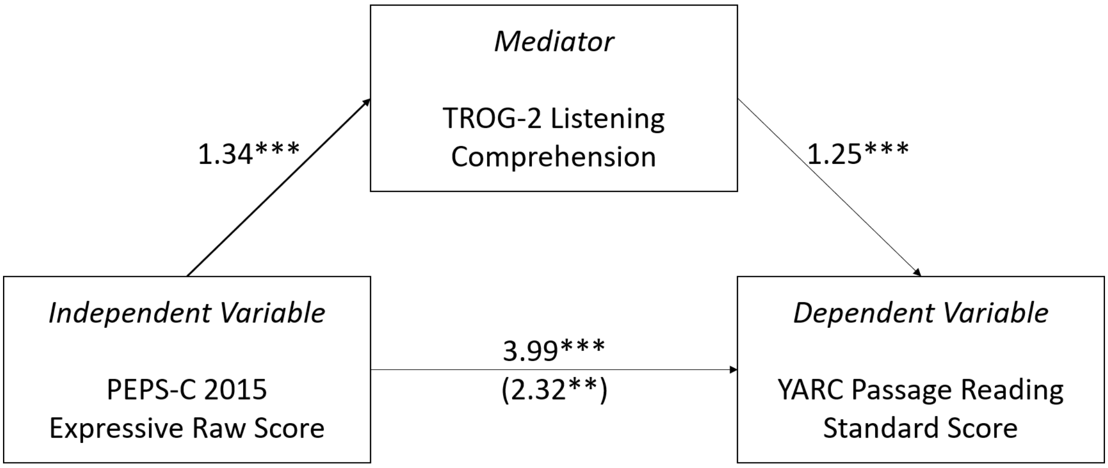

Figure 4.4 Mediation Model 1: The relationship between PEPS-C 2015 Expressive and YARC Passage Reading Standard Score, mediated by the Decoding Composite measure. Note: *p<0.1; **p<0.05; ***p<0.01

The first model created was to look at the effect scores on the PEPS-2015: Expressive tasks had on YARC Reading Passage Standard Scores, when mediated by the TROG-2 Listening Comprehension task. We can see in Figure 4.4 the regression coefficient between the PEPS-C 2015 Expressive score and YARC Passage Reading Standard score was significant both independent of ($b = 3.99$, $p < 0.01$) and when the mediator variable was present ($b = 2.32$, $p < 0.05$), informing us that the presence of TROG-2 Listening Comprehension as a mediator led to partial mediation.

Looking at Table 4.8 we see the output from the mediate function I computed in R. Our model showed a significant effect of listening comprehension on the relationship between expressive prosody and reading comprehension in this model (ACME = 1.67, $p < 0.01$), though despite this there was still a significant direct effect of expressive prosody on reading comprehension (ADE = 2.32, $p < 0.05$).

4.4.2. Mediation Model Two: Receptive Prosody and YARC Passage Reading Mediated by Listening Comprehension

The second mediation model looks at the relationship between PEPS-C 2015 Receptive Prosody scores and YARC Passage Reading Standard Scores, in the presence of the TROG-2 Listening Comprehension mediator. In Figure 4.5 we see the unstandardised coefficients, and see the coefficient between PEPS-C 2015 Receptive and YARC Passage Reading was significant independent of ($b = 3.82$, $p < 0.01$), and when the mediator was present ($b = 2.60$, $p < 0.05$).

Looking at Table 4.8 we see that the there was a significant indirect effect when the mediator was present (ACME = 1.22, $p < 0.01$), which in tandem with the above regressions tells us a partial mediation took place with the presence of the mediator (Proportion mediated = 32%)

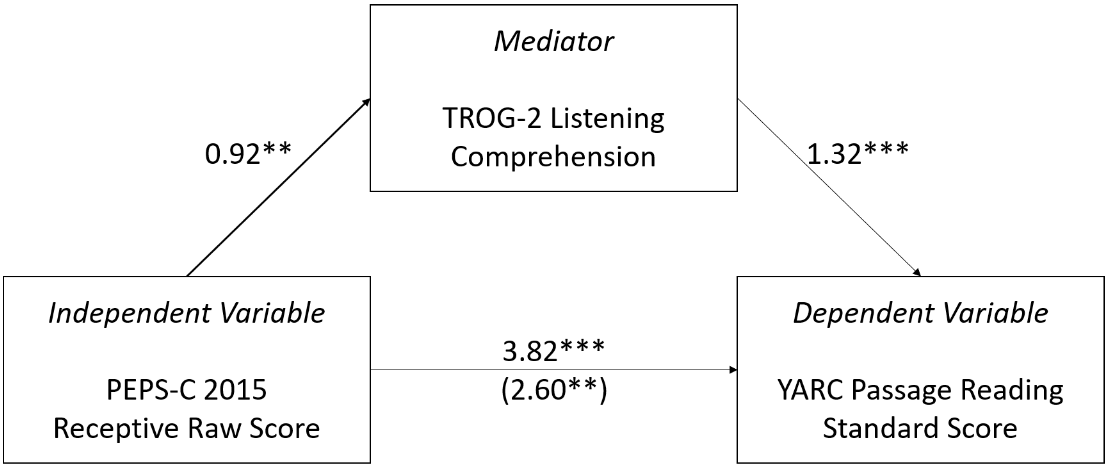

Figure 4.5 Mediation Model 2: The relationship between PEPS-C 2015 Receptive Raw Score and YARC Passage Reading Standard Score, mediated by the TROG-2 Listening Comprehension measure. Note: *p<0.1; **p<0.05; ***p<0.01

4.4.3. Mediation Model Three: Expressive Prosody and YARC Passage Reading Mediated by a Composite Measure of Decoding Ability

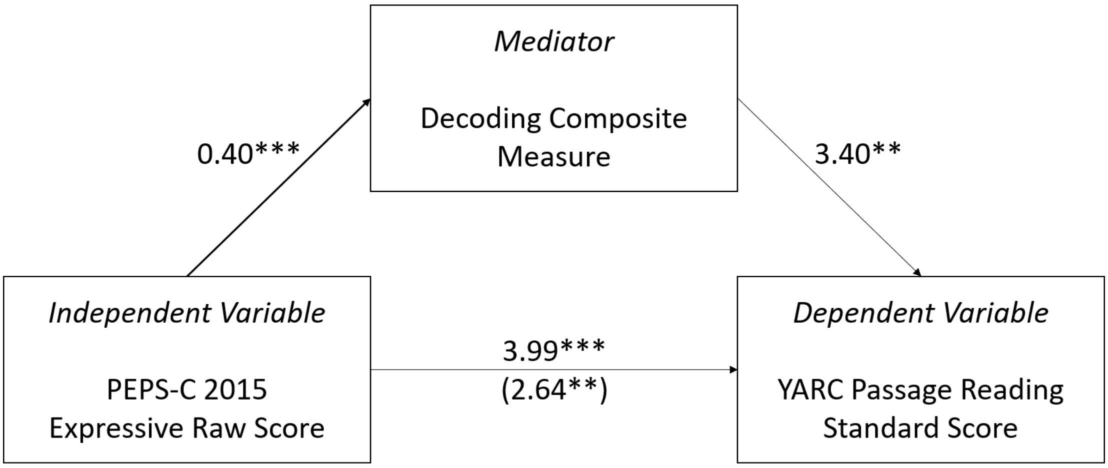

Figure 4.6 Mediation Model 3: The relationship between PEPS-C 2015 Expressive Raw Score and YARC Passage Reading Standard Score, mediated by the Decoding Composite measure. Note: *p<0.1; **p<0.05; ***p<0.01

The third mediation model models the relationship between the PEPS-C 2015 Expressive Prosody scores and YARC Passage Reading Standard Scores, with the Decoding Composite variable as the mediator. The indirect effect when the mediator was present was significant (ACME = 1.35, $p < 0.05$), indicating mediation had occured. Furthermore, the presence of the mediator in the relationship did not stop the direct relationship between the predictor and outcome being significant (ADE = 2.64, $p < 0.05$), telling us that partial mediation had taken place (Proportion Mediated = 34%).

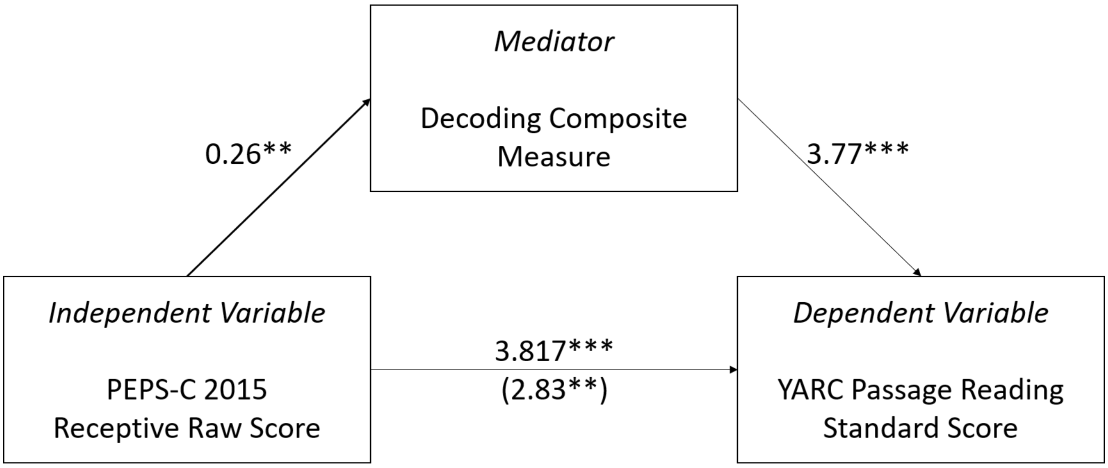

Figure 4.7 Mediation Model 4: The relationship between PEPS-C 2015 Receptive Raw Score and YARC Passage Reading Standard Score, mediated by the Decoding Composite measure. Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

4.4.4. Mediation Model Four: Receptive Prosody and YARC Passage Reading Mediated by a Composite Measure of Decoding Ability

The fourth mediation model models the relationship between the PEPS-C 2015 Receptive Prosody scores and YARC Passage Reading Standard Scores, mediated by the Decoding Composite variable. The indirect effect when the mediator was present was significant (ACME = 0.99, $p < 0.05$), indicating mediation had occured. The presence of the mediator did not stop the direct effect being significant (ADE = 2.83, $p < 0.05$), telling us that partial mediation had occurred (Proportion Mediated = 26%).

Chapter 5. Discussion

The purpose of the current project was to investigate the relationship between children's prosody and reading comprehension, whilst controlling for the two predictive aspects of The Simple View of Reading framework, decoding and linguistic comprehension. I did this by working with 51 year three children in a school in Northern England, and analysing these results afterwards.

There are also very few studies that have looked at the direct relationship between prosody and reading comprehension, so another aim was to add to this base of knowledge. Compared to the other studies in the literature, I used a more systematic and comprehensive approach to measuring prosody with the PEPS-C 2015, which encompassed both of the receptive and expressive measures of prosodic skill. This was following the example of Lochrin et al. (2015), but I created composites of receptive and expressive prosody in an attempt to separate which of these aspects of prosody contributed most to reading comprehension ability. This was the major contribution of the current study, and the hope this choice of method will be taken forward in the future to explore this relationship further.

We now move onto discussing the output from the project in more detail, exploring what our findings tell us about the prosody and reading comprehension relationship, but also how these findings fit into the wider literature. To do this we first discuss the results in relation to the hypotheses and research questions we presented in Section 3.1, to see if our predictions and expectations were met by our results. This is followed by a more general discussion of the results, in which I will relate the output to the prior literature more directly and discuss what we can take from this project, as well as what we may improve or look to investigate in future research.

5.1. Simple Models with Prosody

I begin by discussing the results from the simple models comparing prosody and the other variables. These were conducted to get an idea of how receptive and expressive prosody contributed when not controlling for other measures. It gave me an idea of the relative contribution of receptive and expressive prosody when paired against one another.

5.1.1. Prosody and Reading Comprehension

I begin this discussion by reflecting on the simple relationships between prosody and the other predictors. I predicted that both receptive and expressive prosody would contribute to reading comprehension, based on the work from the literature review (Clin et al., 2009; Kim and Petscher, 2016; Lochrin et al., 2015; Whalley and Hansen, 2006; Whalley, 2017). Of particular interest was the contribution of expressive prosody, as previously stated the majority of the studies I reviewed used receptive prosody as their measurement.

The correlations between the prosody components and the reading comprehension measures were interesting. Expressive prosody showed a strong positive relationship with both the TROG-2 Reading Comprehension ($r = .57$) and YARC Reading Comprehension task ($r = .45$). This makes sense, as both tasks required the reading of sentences or short passages of text, so the need to use appropriate prosody when reading to aid comprehension fits with how I expected the relationship to occur. Receptive prosody on the other hand only showed a significant positive correlation with the YARC Reading Comprehension task ($r = .43$). This may be due to the nature of the task design, the TROG-2 Reading Comprehension required reading out loud, but also required the children to infer meaning from sentences with complicated grammatical constructs. The skills inherent in receptive prosody, that of listening and comprehending prosody in others' speech, did not come into play here, with expressive prosody. Of course, correlation does not imply causation, but the moderately strong correlations added credence to the proposal that understanding and use of expressive prosody aids comprehension of written text.

When modelling both of the prosody components against the measures of reading comprehension I saw more interesting outcomes. For the YARC Passage Reading task both the expressive and receptive components of prosody were significant predictors in the multiple regression model I created. This fits with the above results from the correlation, but it would be good to think why this is. We have already mentioned the expressive prosody contribution, the ability to recognise one's use of prosody and comprehend meaning through this understanding. But, why would receptive prosody also show a significant relationship? The final part of the YARC Passage Reading task required the learners to answer eight comprehension questions, which were read to them by the researcher. This requires the children to listen carefully to the questions asked to help them understand what they need to find in the text to answer the question. As such, receptive prosody is useful in this context as it calls upon the child's ability to monitor and listen to speech, and extract meaning from this. This was not an aspect I had considered when choosing the YARC Passage Reading task, but I feel it is an important potential contributor to receptive prosody's strength in this case. As such, the YARC Passage Reading task may have required both aspects of prosody in different ways.

Comparatively, the TROG-2 Reading Comprehension multiple regression model tells us a different story. In this case, only the expressive prosody measure was a significant predictor of reading comprehension over and above receptive prosody. Thinking about task design again this makes sense as the task required the children to read a grammatically complex sentence out loud and then pick the picture which matched what they had just read. To be successful, it required children to use appropriate prosody and understand what they had just said to pick the correct image. Unlike the YARC task, there was no obvious component of this where receptive prosody skills would come into use, so it seems consistent that this was the outcome. This highlights how important task choice is when designing a project, and an interesting outcome from these regression models.

From these models it seems that expressive prosody was the most consistent predictor of reading comprehension ability, potentially due to the tasks in this composite measure matching the skills required by the reading comprehension tasks most closely. Later in this chapter I will discuss putting these variables in a more complex model with the other variables we measured to see if the relationship maintains with reading comprehension, but for the moment what stood out was expressive prosody being a more consistent predictor of reading comprehension.

5.1.2. Prosody and the Components of The Simple View of Reading

Next I looked at the relationships between prosody and the components of the SVR. As discussed in the literature review, this idea was inspired by Whalley (2017), as I am not aware of any other study that has used the components of the SVR as a control alongside prosody. Before discussing the complex model with all of these variables input, I wanted to look at the contribution of prosody to each of these components.

Beginning with the measures of decoding, a similar story to the previous subsection emerges wherein expressive prosody seems to be the more significant aspect of the two composite measures. For each of our word reading (TOWRE-2 SWE and YARC SWR) and nonword reading (TOWRE-2 PDE) tasks the expressive prosody variable was significantly predictive over and above receptive prosody for each dependent variable. This is potentially due to expressive prosody being most useful to completing these types of tasks. As the children work their way through the decoding tasks, they come to more complex words with more syllables and stress-patterns. As an example, from the YARC SWR word reading task we have the word *endeavour*, which has three syllables, contains complex spelling to decode, and requires stress on the middle syllable. To say this word correctly requires good word knowledge of course, but to get the right stressed syllable would require good knowledge of prosody, as otherwise it would sound very wrong. This highlights where prosody comes into decoding, it aids readers with correct pronunciation alongside just knowing the correct sounds. Receptive prosody was likely not called upon, as this would not help the children with *how* to say the words. Furthermore, in the literature review I highlighted the two projects which had explored multisyllabic word reading in relation to prosody (Holliman et al., 2017b; Wade-Woolley, 2016), which both found a significant contribution by prosody to the reading of multisyllabic words. this fits with the data here, and suggests further work would be viable.

Next is linguistic comprehension, which I expected both elements of prosody could contribute to. The task I used required children to listen to the sentences being read to them, so receptive prosody was likely to be important as this is the skill for the understanding and awareness of prosody use, which is a major aspect of this task. Expressive prosody would also be helpful, for the participants' own awareness of how to produce prosody. The results from the correlation and regression model were interesting, suggesting receptive prosody was not a strong predictor of reading ability as had been expected. Only expressive prosody showed a significant correlation with the linguistic comprehension measure, and as well it was the sole significant predictor in the regression model. This may be due to expressive prosody being a more complex skill, which internalises some of the skill set of receptive skill. I discuss this more in the general discussion section.

The takeaway from these simple models of the SVR is that expressive prosody was the most consistent predictive measure. This suggests in the more complex measures this may be the measure that contributes to reading comprehension more consistently too, but this is what will be explored next.

5.2. Modelling the Relationship between Prosody and Reading Comprehension

I have established that there is a relationship between prosody and reading comprehension, most likely between expressive prosody and reading comprehension. To further explore this relationship I now present the more complex modelling conducted to further assess the relationship. Firstly, I will discuss the mediation models we created. This allows us to see if prosody is mediated by the inclusion of the two predictive components of the SVR, decoding and linguistic comprehension, or if it still produces unique contribution even in the presence of these components. Second, I discuss the two multiple regression models for the reading comprehension measures which included the predictive variables receptive and expressive prosody, decoding, linguistic comprehension, phonological awareness and digit span. By including each of these control measures, which have been shown to contribute to reading comprehension in previous research, we wanted to see if either receptive or expressive prosody would be significant predictors in these particular models.

5.2.1. Mediation Models

I created four mediation models to explore the mediation relationship between the prosody variables and reading comprehension, with the SVR components as the mediators. As mentioned in the previous chapter, we employed this modelling technique as an exploratory measure to think about further research.

Beginning with the models with expressive prosody as the predictor, it was interesting that in both models decoding and linguistic comprehension only partially mediated the relationship to reading comprehension. That means that expressive prosody maintained an independent relationship with the reading comprehension measure. This is interesting to note, and seems to fit with my predictions, that expressive prosody has particular role to play in reading comprehension independent of the SVR framework components. When looking at the receptive prosody predictors mediation models we find the same story, the components of the SVR did not fully mediate the relationship between receptive prosody and reading comprehension.

There is a caveat regarding these models however. There was not enough participants to power the model so we could use multiple mediators and predictors, and this may have changed the outcomes potentially. These models are purely for exploratory purposes, but suggested an interesting relationship that prosody may be a feature of spoken language that could be independent of the components of the SVR. As we discussed in Chapter 1, discussions occur regarding the SVR as to what components are missing and could be identified. There is potential for prosody to be considered here, but more work needs to be done.

5.2.2. Multiple Regression Models

The last models to discuss are the multiple regressions models which contained all the prosody, SVR and control variables in relation to the reading comprehension measures. The first model used the YARC Passage Reading reading comprehension task as the outcome variable, with each of the other tasks regressed onto this. The outcomes showed that only the TROG-2 linguistic comprehension task was a significant predictor for this task at the 0.05 p-value, which was an unexpected finding. I expected linguistic comprehension to have a strong relationship, due to both measures requiring the ability to comprehend language as their aim, just in different modalities, but not for the relationship to be so strong. As well, receptive prosody, though still insignificant, showed a slightly stronger relationship in the model compared to expressive prosody, again unexpected. So what is happening here? The passage reading task was chosen as it was a long-form reading exercise, which tested the children's ability to read for an extended time, then listen to and answer eight comprehension questions. I expected prosody would be significant due to the requirement to read the text out loud (highlighting expressive prosody for comprehending) and listening to the questions to answer the question (receptive prosody for understanding what has been asked of them). This has not been the case, why might this be? My best guess is these skills may have been subsumed by the linguistic comprehension task. I have discussed before that prosody will be an important component of linguistic comprehension, as it helps with comprehending meaning from the suprasegmental aspects of speech. As well, despite how important the reading element of the task was, the children had to listen carefully to the questions at the end of the task to complete the comprehension questions. The results of this task were dependent on answering these questions, so it may be that this task captured more of the linguistic comprehension aspect alongside reading comprehension, as the children could only listen to the questions and not read them. This may have focused the children onto their linguistic comprehension skills, to interpret the question and allow them to return to the text for the answer. As mentioned, prosody was likely a part of this, but linguistic comprehension was likely most important for the outcome of this task. It may be that this task was not as pure a measure of reading comprehension as I thought, or maybe more likely the interlink between linguistic comprehension and reading comprehension is as strong as the SVR framework suggests.

Next is the regression model looking at the TROG-2 reading comprehension task. The same as above, each measure was put in as a control alongside the prosody measures to look at the predictive power on reading comprehension. As noted in the results section, there was an issue with heteroscedasticity in this model which I corrected with a Weighted Least Squares modelling technique. This is likely due to the TROG-2 Reading Comprehension task being a modified version of the original TROG-2 Linguistic Comprehension task, and so had been re-purposed for this project. I chose to do this based on the idea by Whalley (2017) who did the same thing with this task, allowing me to match the linguistic and reading comprehension tasks. Looking at the results of this model it fits more with what I expected, that linguistic comprehension and decoding were significant predictors of reading comprehension, fitting with the SVR framework. But also that expressive prosody was a significant predictor of reading comprehension in this instance. This makes sense thinking about the task, which required the reading of sentences with complex grammatical structure. Expressive prosody likely helped with reading the sentences appropriately with the correct prosodic contour to ensure the meaning came clear to them. The task also just required the children to point at the picture that matched what the sentence said, so it purely relied on their reading of the sentence and ability to comprehend it in the moment. Furthermore, the use of grammatical complicated sentences may be why expressive prosody came into use at this moment. This was an interesting result to come across, and suggested to me that expressive prosody may be an element to explore further in research, being that receptive prosody has been the focus of the work I discussed in the literature review earlier. This gave me an answer to my second research question, alongside the simpler models that expressive prosody was a more consistent predictor of reading comprehension in the current project.

Regarding my first research questions, the answer is mixed. Expressive prosody seemed to provide a unique contribution to reading comprehension for the TROG-2 Reading Comprehension model, and was consistent in the simple measures. But, for the other model linguistic comprehension seemed to subsume this contribution. I feel more confident saying that expressive prosody has the greater potential to be an independent predictor of reading comprehension, as made clear in the simpler models I created. But why might this be? I explored this in the following section, reflecting on expressive prosody and its relationship to reading comprehension as the aspect of my outcomes that has the potential to be researched further.

5.3. General Discussion

This project investigated the relationship between prosody and reading comprehension with children learning to read, whilst controlling for the components of the SVR framework. Prosody is often ignored in studies investigating developmental reading in English, and this project added to the literature exploring prosody's contribution to early reading ability. I also investigated if and how prosody fit into The Simple View of Reading framework (Gough and Tunmer, 1986; Hoover and Gough, 1990), as the framework focuses upon two components that have been seen to lead to successful reading comprehension, but do not include mention of the contribution prosody could have in this relationship. In our discussion of the SVR I mentioned how its focus is upon aspects of speech that involve the visual recognition and understanding of the written word (decoding/word recognition), and the ability to comprehend spoken language which will in turn aid comprehension of written language (linguistic comprehension). Where does prosody fit into this model? That was part of the current project's aim to discover.

In the current project I found using a cross-sectional design with year three children that prosody may be reduced in its independent contribution by the aspects of the SVR. This was seen in the multiple regression models, wherein our listening comprehension measure was the only significant measure of passage reading ability (Table 4.6), though for the TROG-2 Reading Comprehension measure I saw in the OLS and WLS models that our composite measure of expressive prosody was a significant independent contributor, along with the listening comprehension and decoding measures in the WLS version of the model. This was reflected somewhat also in the exploratory mediation models, in which both of the SVR components mediated the direct relationship of the composite measures of prosody with our measures of reading. These results fit with the results of Lochrin et al. (2015), in that they found only an expressive aspect of the prosody tests was a significant predictor of reading comprehension, but was still weaker than other measures in the model. This may be due to the creation of the composite measure of prosody, there were seven individual tasks in each aspect this may have artificially strengthened this relationship to the reading measures through the sheer number of measures.

Regarding the SVR, from the exploratory mediation models both of the predictive aspects of decoding and linguistic comprehension seemed to carry some indirect prosodic influence onto the reading comprehension measures. It may be that prosody is a more important aspect of this framework than first thought, and calls to include it in the SVR are not unfounded. That prosody was not completely mediated suggests that prosody is an aspect that has its own contribution to make to reading comprehension, due to its importance in speech. I discussed in the literature review about prosody being one of the forgotten elements in reading research, and how it could potentially be, even if a small element, an aspect of speech development which aids reading comprehension. However, part of these results could be due to the age group of the current project, being young readers still developing their comprehension ability, and relatively early in their learning to read.

Interesting to note though from the regression models was the insignificant contribution of the phonological awareness measure, the CTOPP-2 Elision task. Phonological awareness is often a significant contributor to reading ability, and is seen as a fundamental aspect of early reading development. There are two possible reasons why it's contribution might have been reduced in the current study. Firstly, I only used one measure to represent phonological awareness, and it may be that this did not capture the skill well enough. I chose the elision task, essentially a phoneme deletion measure, as I felt it would be a challenging measure for the age group who had some experience with phonics in their schooling. It may have been that phonological awareness was subsumed by our decoding measures, which were all word and nonword reading tasks. I discussed that phonological awareness provided a helpful strategy for sounding out new and unusual words, providing children a method to manipulate and understand the words that appear on the page. It may be in the current model that phonological awareness was important, but its contribution was indirectly through the decoding composite measurement, being that phonological awareness is a key to word reading ability. With the prosody measures independent of decoding in the regression models, it is interesting to see that the suprasegmental aspect of phonology may have become more important to the reading relationship when looking at comprehension when compared to the segmental phonology measure.

In Section 2.3 there was discussion around the modelling of the prosody and reading relationship. These were attempts to try and model how prosody fit into speech and reading ability relationship, but were limited by issues such as sample size and the complexity involved in trying to analyse the contributions of these variables between each other. Despite this, the modelling attempts found support that the variables included accounted for significant variance in word reading (Critten et al., 2021; Holliman et al., 2014b). Reflecting back on this model our results fit with some of the pathways explored in these models, particularly that seen in Holliman et al. (2014b) wherein there is an indirect route from prosody to phonemes to word reading. However, I would suggest that we could add a direct pathway from prosody to reading in these models. From the results, expressive prosody was positively correlated with two word reading tasks, the TOWRE-2 PDE ($r = .4$) and the YARC SWR ($r = .45$). Furthermore, with our prosody measures are predictors for the word reading tasks in a multiple regression, expressive prosody was a significant contributor to each individual word reading task. There is a caveat that there is not a regression model including other measures such as the CTOPP-2 or the TROG-2 LC that may have affected this relationship. But, as discussed, when included in the larger regression models for the reading comprehension outcomes prosody was a significant contributor to the TROG-2 Reading Comprehension outcome. Further work may want to engage with recreating these models, as Holliman et al. (2014b) and Critten et al. (2021) did with the Wood et al. (2009) model, and include each of the measures in this model to investigate whether the direct prosody relationship is a reasonable suggestion from these results. In addition, considering the inclusion of both receptive and expressive prosody into the model may be a helpful step to develop this model. For example, expressive prosody may contribute directly to word reading as the regression model in this study, but receptive prosody may contribute indirectly through the phoneme pathway as an example. This adds complexity to the model of course, and the limitations regarding sample size could remain, but there is potential from the current results to expand on these models further.

I end this section thinking about the current project's unique contribution to the reading science field. In attempting to control for the SVR framework I used a well-tested and well-regarded framework of reading ability, and investigated how prosody may fit into this. Our results suggest that prosody (especially expressive prosody) is an important contributor to reading comprehension ability, and may relate to reading through the decoding and linguistic comprehension components. Furthermore, I was able to add to the small set of studies that investigated prosody as measured with behavioural measures against a measure of reading comprehension (Chan et al., 2020; Clin et al., 2009; Deacon et al., 2018; Holliman et al., 2010b; Kim and Petscher, 2016; Lochrin et al., 2015; Whalley and Hansen, 2006; Whalley, 2017) using a multi-component measure of prosody. Prosody was not a significant predictor for the YARC Passage Reading task, which may be due to the nature of this particular task. It required reading an extended piece of text and monitoring it to answer comprehension questions after. This may be why the measure of linguistic comprehension was the only significant predictor, as similar skills were being used to try and complete the task. But, expressive prosody component was a more consistent predictor of reading comprehension across the different analyses conducted, including the correlation measurements and in the models containing only receptive and expressive prosody. This was the unique aspect of my study, that expressive prosody may be predictive of reading comprehension above receptive prosody, and to investigate prosody's further researchers should use more expressive prosody tasks in their modelling. All of the studies I mentioned in the literature review used receptive prosody tasks, but only Lochrin et al. (2015) highlighted the use of expressive prosody tasks. This is likely a future avenue for research.

But, why was the expressive prosody component more consistent in its predictive strength? Part of this may be down to expressive prosody tasks requiring greater prosody skill than receptive tasks. To be successful as a receptive task, you might be able to recognise the differences in prosody, but not be able to use prosody effectively yourself. Whereas, for expressive prosody you would expect the speaker to both recognise differences in prosody as well as use prosody effectively to be able to produce meaning. This means expressive prosody tasks may well be more sensitive to individual differences in prosody use and understanding. This may be why expressive prosody was significant over and above receptive prosody in all of the models against the SVR components, it is more sensitive to prosody skill as it is a more complex aspect of prosody to develop.

I suggest future studies utilise multi-component measures of prosody, maybe investigating the individual tasks rather than as composites, to try and disentangle what aspects of prosodic skill relate to differing aspects of reading ability. I used composite representations of the prosody tasks in the current project as I wanted to try and pry apart the receptive and expressive relationship, and dividing the individual tasks into these composites was the most sensible way to achieve this. With this in mind, I would suggest that focus could be moved to focus more on the expressive aspect of prosody at the earlier stages of learning to read, including more behavioural measures allowing children to show their understanding of how they use prosody in their own spoken language.

5.3.1. Limitations

Number of Participants

The first major aspect was the number of participants I was able to get for the study. The 51 pupils I worked with whom fit the criteria I set in the Section 3.4 are a small sample of childhood development. Furthermore, in working with only one year group (year three) the sample was reduced to a very specific group, which hinders generalisability. I made concessions to aid this, using bootstrapping procedures to re-sample the data and analyse the confidence intervals this produced to try and account for this issue. This can only take us so far, and the sample size meant I did not have the power necessary to conduct the multiple mediation model I had originally planned, allowing us to control for both decoding and linguistic comprehension as mediators in a single model. As such, I conducted multiple simple mediation models, parsing out the components of the SVR as separate mediators, and the PEPS-C 2015 Expressive and Receptive tasks as separable predictors. This is not ideal, and calls to the issue of conducting multiple measures and making comparable analyses. Due to the COVID-19 pandemic our plans to recruit more participants was stymied for follow-up work, so I was unable to work with more children and this has ramifications for the outcome of the study. In conducting a cross-sectional study with one age group there are already issues with an inability to investigate causation in the relationships between the variables I measured, however I feel that in the scope I was able to complete the exploratory mediation measure was suitable for the current study. It gave me an idea of what could be happening with these relationships, and provides ideas for future research to explore.

Cross-Sectional Design

The project was designed to measure speech and reading ability at one point in time, with a single age group of children. I succeeded in that case, but the issue is that I only have one time point and a single age group of children to discuss. This weakened or prevented us from discussing aspects of causation or longer term implications of the relationship between prosody and reading. I will discuss this in the following section, but to better understand the prosody and reading relationship longitudinal studies have to be conducted, to see how prosody shifts in strength and contribution. With this current study, I learnt that expressive prosody skills are a potential predictor of reading comprehension ability, and is an exciting development that can be taken in future methodology designs.

Linguistic Comprehension Measurement

Another limitation that became clear following data collection was in regards to linguistic comprehension. As discussed in Subsection 1.1.1 in discussion of the SVR it was mentioned that from the literature that researchers should aim to match the method used to measure linguistic and reading comprehension. This is because there is a thought that linguistic and reading comprehension call upon the same higher order cognitive skills, and differ just in the medium in which they are presented (this is a view which can be fairly critiqued, but for the current study I followed this idea as I wanted to explore the SVR). For the TROG-2 task, I was inspired by Whalley (2017) and their modification of this task to create a matched listening comprehension and reading comprehension version using the same stimuli but swapping the nouns, along with counterbalancing the tasks by swapping which one each child encountered first. The issue is that I did not conduct the same matching of methodology with the YARC Reading Comprehension Test task. I could have taken passages of the same level from the YARC stimuli and adapted them to a listening comprehension task, simply by reading the tasks and asking the questions afterwards. Furthermore, these would have been narrative and expository tasks matched by standardised level too. I only came to this realisation following data collection. This was an oversight, and further research should match these tasks when using the SVR as their framework.

Measuring Prosody Solely with Behavioural Measures

A decision was made regarding measuring prosody, that the measurement would be of receptive and expressive prosody recorded through verbal and non-verbal means, but with no recording and analysis of the speech stream. This meant I did not analyse the use of prosody in speech, like that conducted in Benjamin and Schwanenflugel (2010). It was beyond the scope of the current project, but future research could take the methodology in this study and investigate individual differences in prosody use and focus on children's ability to focus on the specifics of prosodic expression (e.g. timing, rhythm).

No Measure of Morphological Awareness as a Control Measure

As is clear from Chapter 1 and the discussion of prosody and reading comprehension, morphological awareness is a potentially important control measure to include when investigating this relationship. Through each of the studies by Deacon et al. (2018), Chan et al. (2020) and Clin et al. (2009) it seemed that prosody's contribution was either reduced or eliminated through the presence of such a measure. The design for the current project had been completed before the publication of the studies by Deacon et al. (2018) and Chan et al. (2020), and so the full significance of morphological awareness was not fully appreciated based on the basis of the study by Clin et al. (2009) alone. There is a possibility that with morphological awareness that prosody would not be related to reading comprehension at all, and this is an interesting area of research that can still be explored further.

5.3.2. Future Directions

Through the completion of the project, there were many ideas I had which I could not include. Every project can be better, and in this section I highlight some ideas that future researchers could and should take forward to better understand prosody and its relationship to early reading development.

Prosody and Morphological Awareness

Based on the research discussed in Chapter 1, this appears to be an important aspect to explore. There is little to add beyond what was discussed in the Limitations section, only to say that future studies should plan to include a measure of morphological awareness as a control measure, to see if its presence changes prosody's relationship with reading comprehension.

Matching Linguistic and Reading Comprehension Measures

Also mentioned in the previous subsection, but it is worth repeating. There is good reason to try and match measures theoretically, the justification being that you hone in on the difference between these skills. The only difference being the form of presentation, which helps to emphasise listening or reading skills. In the current study, this was attempting with the TROG-2 task. There was the version of the task where the students listened to the researcher read the sentence for comprehension, and a version where they read the sentence on the page out loud. Both tested their comprehension ability, but altering the method the student interacted with the sentence to comprehend helped to highlight a specific skill.

Further work could concentrate on matching all of the linguistic and reading comprehension measures in the SVR framework. Under this framework the only difference in descriptions between these two aspects is how the task is presented. In the current project there was the YARC Reading Comprehension passage reading task, this could have adapted this task to be a listening comprehension task where the child listened to the researcher read a piece of text. Once they heard the narrative, the children could then answer the eight comprehension questions, testing their retention of the information from their listening.

Eye-Tracking as an Extra Measure of Reading Ability

An area of research that has been gaining traction in child language research is the use of eye-tracking. Eye-tracking is the recording of movement of a participant's eyes, utilising the reflection of the pupil and cornea to estimate gaze (Holmqvist et al., 2011, p. 24). The thought is that, over and above traditional measures of reading used in research (such as in this current project), tracking eye movements provides us an insight into moment-to-moment cognitive processing during the course of reading (Kaakinen et al., 2015). In conjunction with behavioural methods, it may yield more sensitive and precise data about the behaviour during reading (Ashby et al., 2013). The thought is that this method can allow researchers to investigate the process of reading in a naturalistic state for the participant. This multimodal approach to reading ability measurement seems to be good for investigating reading processing as it happens (Ashby et al., 2013).

Eye-tracking research with children investigating reading is still a work-in-progress, with a review by Blythe (2014) identifying fewer than 30 studies which had included a sample of children looking at eye movements related to developmental reading. Why might this be? It is because conducting eye-tracking studies with young children is difficult, as even with more user-friendly and a move to video-based equipment children have to be willing to cooperate for the extended duration of a full study. This involves sitting very still whilst their eyes are calibrated, following instructions and remaining patient. It's tough enough conducting eye-tracking studies with adults, adding younger participants into the mix is another layer of difficulty.

This previous paragraph is not to put individuals off using eye-tracking, rather to know that using this method with children requires careful implementation and serious thought. But, there is a dearth of eye-tracking studies comparing developing readers against highly-skilled readers to better understand the reading process (Blythe and Joseph 2011, p. 643; Miller and O'Donnell 2013; Rayner et al. 2013; Schroeder et al. 2015). In reviewing the research, Blythe (2014) found that there was general trends for the control of eye movements to get better with age and development of reading skills, as if eye movement behaviours mature to become more efficient over time in reading.

Further to this, I discussed implicit prosody in Section 2.8.2, and there is research conducted exploring the Implicit Prosody Hypothesis. One example is the use of garden-path sentences (e.g. We painted the wall with cracks), examining how what eye movements occur when an individual has to re-read and reassess what they have just read to make sense of it. If unable to do a task with eye-tracking technology itself, use of methods such as self-paced reading offer an alternative manner to try and capture this implicit prosody skill. Self-paced reading is programmed so that words or parts of text appear on screen a bit at a time, and the reader controls when the next piece of text appears by pressing a button. What this allows is to replicate the eye movements readers make from word-to-word, and measure how long children focus on the section of text you display before moving to the next section. Furthermore, garden-path sentence design can be used here to lead children to an assumed reading of the sentence, but then surprise them through the use of grammar as to what the sentence is about.

Eye-tracking was originally intended for use within this thesis, but due to budgetary concerns getting the right eye-tracker to take into schools, and the onset of the COVID-19 pandemic this plan had to be rejected. Future researchers should think seriously about using eye-tracking as a cognitive measure, in relation to the more common behavioural measures used in reading. Behavioural measures are of great utility, but reading is a physical skill too through the control of eye movements and a multi-method approach to this could give us greater insight in the development of reading ability in the early years.

Design Project Across Year Groups and Longitudinally

Due to the time and availability of participants, the current project was only able to work with a single school year group for data. This meant our results were tempered somewhat and taken as exploratory for the more complex models I created. Individuals intending to investigate the relationship between prosody and reading comprehension further would be better comparing multiple age groups with a cross-sectional design. By comparing multiple age groups one will get a better idea of the shifting relationships to reading comprehension. For example, I discussed in the context of the SVR that the relationship between decoding and reading comprehension changes over time, wherein the ability to decode written language is one of the initial reading skills children in UK schools develop through the phonics programme. Once developed, learning to read changes to reading to learn, which is where linguistic comprehension may become a stronger component of the SVR framework's equation.

The reality of conducting longitudinal studies is not as simple as just following a group for years, it requires much time and resources, which is why it is not the best design for a thesis project unless one comes into the project knowing they will be doing this from the beginning. However, researchers who have the necessary skills and resources to tackle such a project would add a lot to the field with such a design. To track and measure developing reading ability through the early education of children is an undergoing only established and skilled researchers can tackle, but if we are to understand the trajectory of reading ability throughout childhood such studies are important to attempt.

5.3.3. Closing Statements

The current project investigated if prosody contributed to reading comprehension through the framework of The Simple View of Reading. Prosody did indeed show some significant contributions to reading comprehension, with both receptive and expressive prosody showing strong relationships with the reading comprehension tasks I chose. It may be that these two aspects of prosody could be focused on more closely in future studies, as I am only aware of this study and Lochrin et al. (2015) which used these two definitions in the literature. There were limitations related to sample size and cross-sectional design which can be improved upon on in future research. This is why in the Future Directions section I wanted to express that future researchers plan carefully with child samples over a longer period of time, which is hard to achieve, but to understand how prosodic ability contributes to reading this is likely the only way to achieve this.

In closing, the complexity of prosody is a continuous aspect of this research, how one chooses to define prosody affects so much of the measures you choose, the design of the project and more.

The current project identified prosody as an aspect of language which can be contributory to success in reading comprehension, and that further research could investigate more about the individual contributions of receptive and expressive prosody. There were many examples of bespoke prosody tasks in the literature review which looked at receptive prosody, there is a great gap and opportunity to create new measures exploring this and I hope other researchers explore this area further. Furthermore, I feel the current project identified expressive prosody as an avenue for designing new prosody tasks which could help with understanding the prosody and reading comprehension relationship. Overall, there is still work to be done exploring the contribution of prosody to reading comprehension, and I hope this piece of research provides some direction forward.

Appendix A. Ethics Forms

A.1. Parent's Consent Form

Consent Form for Volunteers (Parents)

School of English Literature, Language and Linguistics,
Percy Building,
Newcastle University,
NE1 7RU

CONSENT FORM FOR VOLUNTEERS

(Parents)

Your child has been invited to take part in a research study. Before you decide whether they can take part, it is important for you to understand what participation will entail.

Please take your time and read the information sheet provided by the researcher and ask any questions you may have before agreeing to have your child take part. After completing this document send it back into school with your child to be collected by the researcher.

PROJECT TITLE

Speaking to Reading in English: Investigating the Relationship between Prosody and Early Reading Comprehension with Year 3 Children

---

RESEARCH WORKERS, SCHOOL AND SUBJECT AREA RESPONSIBLE

Sheradan Miller (PhD Student and Lead Researcher)

Dr. Cristina Dye (Supervisor and Project Supervisor)

School of English Literature, Language and Linguistics, Newcastle University

---

If you agree to let your child take part in this experiment, they will be asked to do the following:

- Complete their own consent form to ensure they wish to take part in the study.
- You and your child will be given the opportunity to ask questions or voice any concerns about the study before it begins. This can also be done any time through the study.

CONFIDENTIALITY OF INFORMATION

The confidentiality of personal information and the anonymity of all volunteers involved in this investigation will be preserved in the following way:

- All participants will be assigned a personal code to identify them only for data processing purposes. Yours and your child's name will be recorded on the corresponding consent forms only; any discussion of results arising will be anonymous.
- Recordings of children's speech will only be heard by the lead researcher for data-checking, these recordings will be erased at the end of the project.

1

Figure A.1 Parent's Consent Form Page 1

Consent Form for Volunteers (Parents)

- Your child's data will not be shared with anybody with potentially identifying information. Data will be fully anonymised before being shared with others for discussion and analysis.
- Any questions about your data, the experiment, or anything else can be directed to the lead researcher, who will be happy to discuss any enquires.

CONTACT INFORMATION

If you wish to withdraw your child's data from the study, please contact Sheradan Miller (*s.miller2@newcastle.ac.uk*) within 28 days following the end of the experiment.

If you wish to make a complaint about the way the research is conducted, please contact the Project Supervisor in the first instance: *cristina.dye@newcastle.ac.uk.*

VOLUNTEERS STATEMENT

Please read and think about the five points below and ensure you agree with each statement before confirming you are happy for your child to take part in the research:

1. I confirm that I have read and understood the information sheet for the proposed project and have asked any questions I wanted to.
2. I understand that my child's participation in this research is voluntary and they are free to withdraw at any time, without giving a reason.
3. I understand that all information collected
    a. Will be labelled with a unique code representing my child, and not my child's name to maintain anonymity.
    b. Will be stored in a secure place only accessible by the lead researcher and their project supervisors.
4. I understand that I can contact either the lead researcher or project supervisor with any queries I have about the research following the study.
5. I understand that if my child discloses information that they may be at risk it will be followed up through established channels.

If you agree with each of the points above, please complete and sign the form below.

Signed by Parent/Guardian

Name of Child:

Name of Parent/Guardian:

Signed:

Dated:

Signed by Researcher

Name:

Signed:

Dated:

2

Figure A.1 Parent's Consent Form Page 2

A.2. Parent's Information Sheet

School of English Literature, Language and Linguistics,
Percy Building,
Newcastle University,
NE1 7RU

Information Sheet for Parents/Guardians

Project Title: Speaking to Reading in English: Investigating the Relationship between Oral Language and Early Reading in Children

Dear Parent/Guardian,

Your school has agreed to take part in a research project investigating the relationship between oral language and reading ability in young children. Your child's class has been asked to take part, and this information sheet is to inform you about the projects aims and answer some important questions we feel you may have.

What is the project about?

The purpose of the project is the explore the relationship between young readers' oral language skills and early reading ability. Children will be asked to complete various tasks involving paper and pencil games, listening games and computer games to help us understand the relationship between oral language and reading. One example is the affect task, in which we will present sentences to the children where a person will say a food item ('mushrooms') in a accepting or disapproving tone, and the child identifies whether the voice likes or dislikes the food.

To complete all the tasks should take around 2 hours over several shorter sessions, and your child will be provided with breaks to ensure they don't get tired completing the tasks.

Where will the research take place?

The research will take place within your child's school during the regular school timetable. They will work with the researcher in a pre-determined space chosen by the school for 2-3 short sessions of around 30-40 minutes each.

Why do researchers want to work with my child?

Young children are amazing in that they are in the process of developing all the skills they will have for the rest of their life. It is great for researchers to capture data at these early stages of development as it provides us insight into how and why children learn as they do. In reading research, we particularly want to understand more about the early stages of learning to read and how it develops, so we want to try and create engaging research for children to complete which can help more children with reading in the future.

Why would I want my child to participate?

We hope that by allowing your child to participate that they can be a crucial part of influencing cutting-edge research attempting to help children become better readers through understanding how children learn to read. As part of the project, you will receive a debrief form explaining what the project is about in greater detail, and be asked if you want to receive updates on the project as it continues to develop. As part of their involvement we

Figure A.2 Parent's Information Sheet Page 1 also want to reward your child, so will provide them with a certificate of participation and a sticker card tracking their progress through the tasks.

Why would my child want to participate?

We have designed the tasks to be more like a game than a test, so we hope the experience will be enjoyable and fun for the child as we collect data about them. Furthermore, the tasks are designed to be age-appropriate and suited to their interests. Children are also not judged as “right” or “wrong” during the session, we simply want to see how they respond to the tasks and will provide positive affirmation as they help us. Your child is also able to request to the stop the session at any time. We find children enjoy being part of research and come away with a positive attitude about what they have been a part of.

What ethical and security procedures are in place?

All projects conducted at Newcastle University must go through full ethical approval from the Newcastle University Ethics Committee, which this project has. Furthermore, this project is a PhD project and has had to go through a project approval phase to ensure it is suitable for conducting with Newcastle University’s permission. The lead researcher who will work with your child has an up to date DBS check and has prior experience working with children in a research setting.

For certain tasks children’s vocal responses will be recorded for later data-checking to ensure results are accurate. These recordings will be anonymised and accessible only by the lead researcher, and will be deleted following the conclusion of the project.

You and your child will each provide consent to take part in the project, and no data will be collected without this consent. Furthermore, your child’s data will be stored digitally within the University for viewing and analysis by the lead researcher and their supervisory team. A backup of this data will also be stored on a password protected external hard drive also kept within the University. Your child’s data will be kept anonymised always, represented only by a participant number with no personally identifying information kept with this data.

What do I do next?

If you want your child to take part in the project firstly thank you, secondly please sign the consent form supplied with this information sheet and ask your child to return it to their class teacher. The researcher will then come to school and meet with your child to sign their own consent form and conduct the research. Children will be taken out of class **ONLY** if a signed consent form has been returned by their parent.

How do I get in contact if I have any queries?

If you wish to inquire about more information regarding the study, please contact lead researcher Sheradan Miller (s.miller2@newcastle.ac.uk) or his supervisor Dr Cristina Dye (cristina.dye@newcastle.ac.uk).

Figure A.2 Parent’s Information Sheet Page 2

A.3. Parent Questionnaire

School of English Literature, Language and Linguistics,
Percy Building,
Newcastle University,
NE1 7RU

Parental Questionnaire

Thank you for taking time to register interest in the project. Before we work with your child within the school, we kindly request you complete the questionnaire questions written below. This data will not be used for anything but identifying whether your child is suitable for the project and for demographic information. The sheet will be separated from the consent once the lead researcher has logged the information digitally to ensure confidentiality.

1. What is your child's date of birth?

   ____

2. Is English your child's first language?

   Yes ☐ No ☐

3. Which hand does your child use predominately?

   Left ☐ Right ☐

4. Does your child have any learning difficulties that you are aware of? (e.g. dyslexia)

   Yes ☐ No ☐

5. Does your child have any neurological difficulties that you are aware of?

   Yes ☐ No ☐

Thank you for taking the time to complete this questionnaire. Please send it back to school with your child along with the completed consent form.

<u>CONTACT INFORMATION</u>

If you wish to inquire about more information regarding the study, please contact lead researcher Sheradan Miller (s.miller2@newcastle.ac.uk) or his supervisor Dr Cristina Dye (cristina.dye@newcastle.ac.uk).

Figure A.3 Parent Questionnaire

A.4. Parent Cover Sheet

School of English Literature, Language and Linguistics,
Percy Building,
Newcastle University,
NE1 7RU

Dear Parent/Guardian,

My name is Sheradan Miller, and I am a PhD researcher at Newcastle University. I am contacting you as I am currently conducting a project investigating how children's proficiency with speech helps them to become better readers. As part of this, I am going to be coming to Fulwell Junior School to work with children in Year 3 through the final half term before the Summer break. This will involve participating in games involving word reading, listening and speaking, and reading comprehension. See the attached information sheet for more information.

For your child to take part we need permission from a parent/guardian. Please read the attached documentation to help you understand what the project is about. Following this you will need to complete the **Consent Form** and **Parent Questionnaire** for your child to be able to take part. Please return the consent form and parent questionnaire with your child to give to their class teacher, and then I will be able to work with them in school.

If want to know more information about the study, you can contact myself (s.miller2@newcastle.ac.uk) or my supervisor Dr Cristina Dye (cristina.dye@newcastle.ac.uk), and we will be happy to answer any of your questions.

Thank you for taking the time to read this and I hope you I get to work with your child for the project.

Yours faithfully,

Sheradan Miller (2nd Year PhD Student)

Figure A.4 Parent's Cover Sheet

A.5. Children's Consent Form

Consent Form for Volunteers (Child)

Newcastle University

School of English Literature, Language and Linguistics,
Percy Building,
Newcastle University,
NE1 7RU

Child's Code:

CONSENT FORM FOR VOLUNTEERS

(Child)

Thank you for letting me visit you in school today to take part in some reading games!

Before we start I want to tell you what we will be doing today. Then, I will let you ask me any questions you have about what we are going to do.

We are going to play some games using paper and a computer. You will be listening to sounds, reading words and sentences, and some other games. I want you to answer them as best you can, don't worry if you don't know just try your best!

Then we will use the eye-tracker. I will help you get comfy and take you through the game. You will just be reading for me and I will do the rest.

I will also use a voice recorder to record you speaking for some games. No one else will hear this apart from me. I will use it to check what I have written down when you were speaking to me is correct.

If you want to stop playing the games at any time you just need to tell me, and we will stop. Also, if you want to ask me any questions then just ask me and I will answer them for you.

Please circle the smiley face telling me whether you want to play the games or not below.

YES **NO**

Thank you for answering!

Please write your name below and we will get started!

Name:

Date:

Researcher's name:

Researcher's signature:

Figure A.5 Children's Consent Form

Appendix B. Measures

B.1. Comprehensive Test of Phonological Processing - Second Edition (CTOPP-2): Elision

CTOPP-2 Elision Task Stimulus Sheet

Participant Number:

Date:

"Let's play a word game."

1. Say "toothbrush". Now say "toothbrush" without saying "tooth" - **BRUSH** [1 / 0]
2. Say "cowgirl". Now say "cowgirl" without saying "girl" – **COW** [1 / 0]
3. Say "popcorn". Now say "popcorn" without saying "corn" – **POP** [1 / 0]
4. Say "baseball". Now say "baseball" without saying "base" – **BALL** [1 / 0]
5. Say "sunshine". Now say "sunshine" without saying "sun" – **SHINE** [1 / 0]
6. Say "airplane". Now say "airplane" without saying "plane" – **AIR** [1 / 0]
7. Say "always". Now say "always" without saying "all" – **WAYS** [1 / 0]
8. Say "doughnut". Now say "doughnut" without saying "dough" – **NUT** [1 / 0]
9. Say "spider". Now say "spider" without saying "der" – **SPY** [1 / 0]

"Okay, now let's try some where we take away smaller parts of the words"

10. Say "cup". Now say "cup" without saying /k/ - **UP** [1 / 0]
11. Say "meet". Now say "meet" without saying /t/ - **ME** [1 / 0]
12. Say "farm". Now say "farm" without saying /f/ - **ARM** [1 / 0]
13. Say "mat". Now say "mat" without saying /m/ - **AT** [1 / 0]
14. Say "bold". Now say "bold" without saying /b/ - **OLD** [1 / 0]

DO NOT GIVE FEEDBACK IF INCORRECT ANSWER GIVEN FROM THIS POINT

15. Say "tan". Now say "tan" without saying /t/ - **AN** [1 / 0]
16. Say "time". Now say "time" without saying /m/ - **TIE** [1 / 0]
17. Say "mike". Now say "mike" without saying /k/ - **MY** [1 / 0]
18. Say "snail". Now say "snail" without saying /n/ - **SAIL** [1 / 0]
19. Say "sling". Now say "sling" without saying /l/ - **SING** [1 / 0]
20. Say "winter". Now say "winter" without saying /t/ - **WINNER** [1 / 0]
21. Say "powder". Now say "powder" without saying /d/ - **POWER** [1 / 0]
22. Say "faster". Now say "faster" without saying /s/ - **FATTER** [1 / 0]
23. Say "silk". Now say "silk" without saying /l/ - **SICK** [1 / 0]
24. Say "driver". Now say "driver" without saying /v/ - **DRYER** [1 / 0]
25. Say "tiger". Now say "tiger" without saying /g/ - **TIRE** [1 / 0]
26. Say "flame". Now say "flame" without saying /f/ - **LAME** [1 / 0]
27. Say "strain". Now say "strain" without saying /r/ - **STAIN** [1 / 0]
28. Say "splat". Now say "splat" without saying /l/ - **SPAT** [1 / 0]
29. Say "planes". Now say "planes" without saying /n/ - **PLAYS** [1 / 0]
30. Say "split". Now say "split" without saying /p/ - **SLIT** [1 / 0]
31. Say "stride". Now say "stride" without saying /s/ - **TRIED** [1 / 0]
32. Say "banks". Now say "banks" without saying /k/ - **BANGS** [1 / 0]
33. Say "pixel". Now say "pixel" without saying /s/ - **PICKLE** [1 / 0]
34. Say "fixed". Now say "fixed" without saying /k/ - **FIST** [1 / 0]

Figure B.1 CTOPP-2 Elision Stimulus Sheet: These were the instructions the researcher used to conduct the Elision task. Wording was following exactly the same as on the sheet.

B.2. Working Memory Test Battery for Children (WMTB-C) - Backwards Digit Recall

Working Memory Test Battery for Children (WMTB-C) - Backwards Digit Recall Stimuli

Practice Stimuli
2-3

5-4

3-4-5

5-2-4

Two Numbers
2-7

5-9

3-1

9-7

4-6

8-4

Three Numbers
8-1-4

6-3-7

4-6-2

2-5-9

7-3-5

9-4-3

Four Numbers
2-7-1-4

5-2-7-3

6-3-8-4

1-5-4-9

9-6-5-8

8-1-6-2

Five Numbers
2-1-4-9-8

5-7-1-4-2

2-7-4-6-3

9-5-1-4-2

3-5-8-2-6

4-6-3-1-5

Six Numbers
5-2-1-7-9-3

2-7-6-3-8-5

4-8-3-5-2-7

8-5-2-9-1-3

1-9-5-8-2-4

6-1-3-9-5-2

Seven Numbers
8-3-5-2-9-4-1

6-3-1-9-4-7-5

5-8-7-2-4-9-3

7-9-2-6-1-9-3

8-5-2-4-9-3-6

9-6-2-8-1-4-7

Figure B.2 WMTB-C Backwards Digit Recall: Children had to successfully recall each of the practice stimuli backwards, then they would move to the actual stimuli beginning with the “Two Numbers" section. Children had to recall at least 4 of 6 of the stimuli correctly to move to the next section.

B.3. Profiling Elements of Prosody in Speech-Communication (PEPS-C 2015)

B.3.1. Familiarisation with Vocabulary

The figures below are examples of the screens participants saw when conducting the familiarisation task.

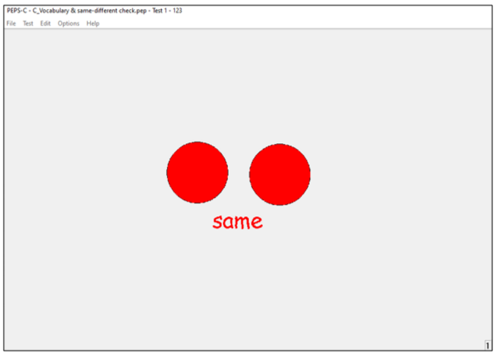

Figure B.3 Familiarity check for “same".

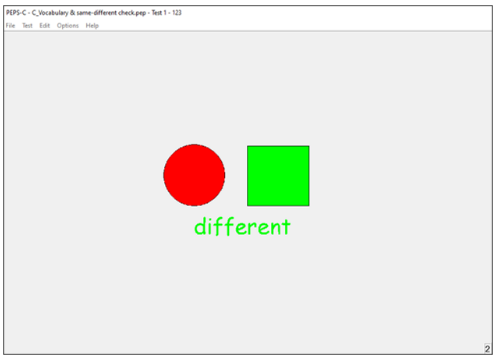

Figure B.4 Familiarity check for “different".

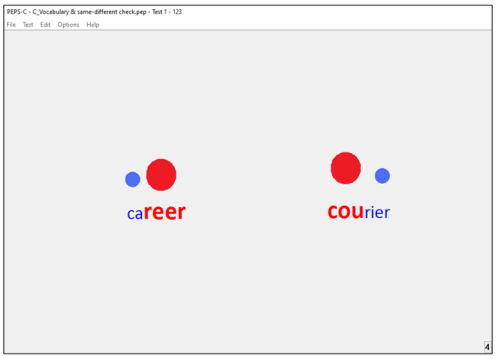

Figure B.5 Familiarity check for stress differences.

Figure B.6 Familiarity check for stress boundaries.

Figure B.7 Familiarity for vocabulary used in tasks.

B.3.2. *Receptive Short Item Discrimination*

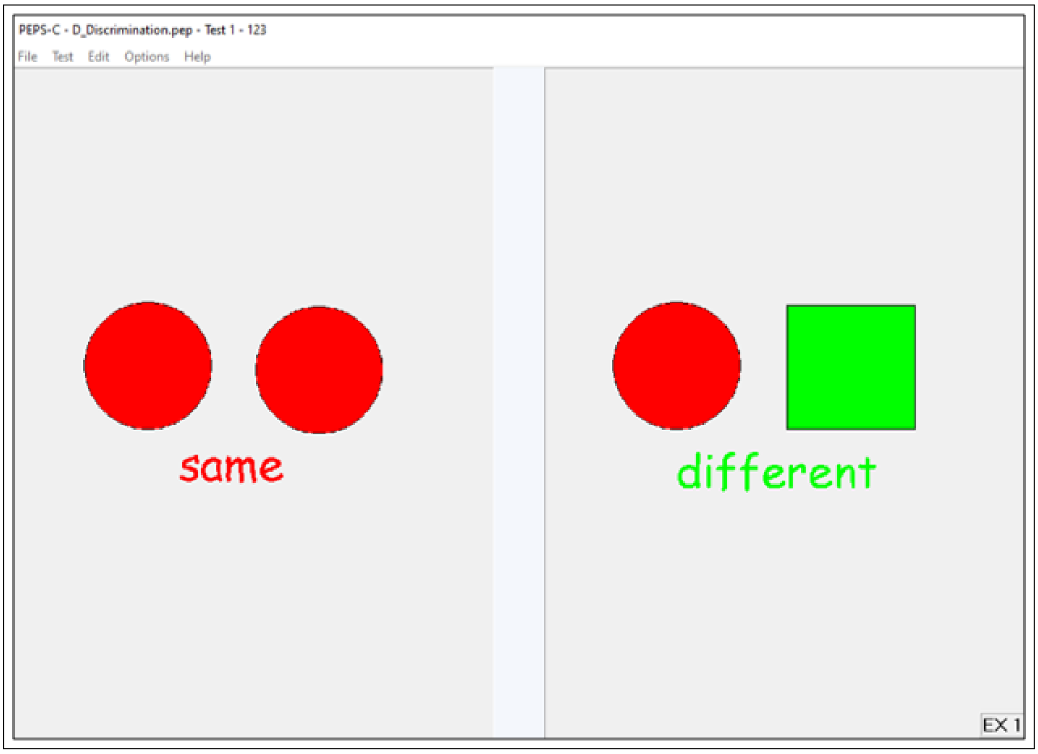

Figure B.8 PEPS-C 2015 Auditory Discrimination Task: This task tested children's ability to perceive prosodic differences in short utterances. The utterances were edited to remove phonological information, leaving only the prosodic contours for comprehension. Children heard two utterances one after the other, and were instructed to listen carefully to identify whether the utterances sounded the same or different. To log their answer, the children either clicked on the word 'same' or the word 'different' on the computer.

B.3.3. Expressive Short Item Imitation

Figure B.9 PEPS-C 2015 Imitation Task: This task tested children's ability to listen to and imitate different forms of intonation with words or phrases. Children listened to the voice recoding on the computer say a word. The children were then asked to imitate the voice as close to the intonation of the speaker. The stimuli were items the child had been made familiar with. The researcher scored the children using the keypad; they pressed G for a good imitation, they pressed A for an average imitation, and finally they pressed P for a poor imitation.

B.3.4. Receptive Turn-end

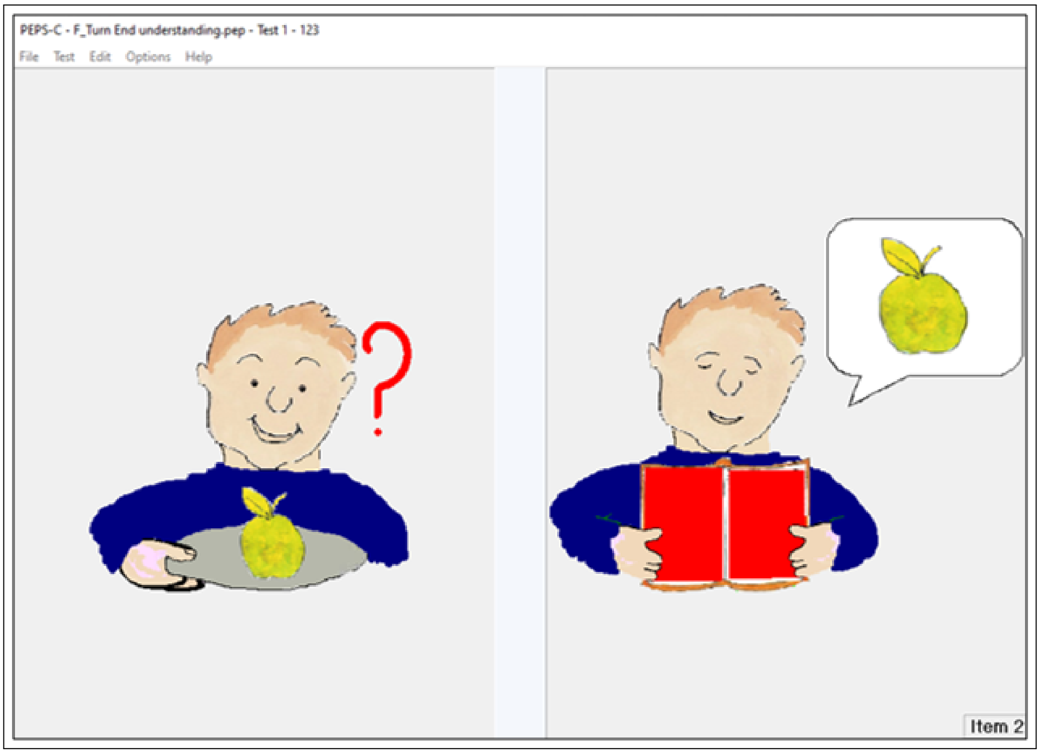

Figure B.10 PEPS-C 2015 Receptive Turn-end Task: This task tested whether children could perceive the difference between questioning and declarative intonation. Two pictures appeared on the screen (see above) representing these intonations. The child then listened to the voice recording and decided whether the food item was being said with a questioning or declarative intonation.

B.3.5. Expressive Turn-End

Figure B.11 PEPS-C 2015 Expressive Turn-End Task: This task involved children being shown the same pictures from the task in Appendix B.3.4, but this time they were required to say the food item with the appropriate intonation denoted by the picture. The researcher used the keypad to identify whether the vocal intonation was questioning (press ?), declarative (press ✓) or ambiguous as to the intonation used (press A).

B.3.6. Receptive Affect

Figure B.12 PEPS-C 2015 Receptive Affect Task: This task involved children listening a pre-recorded voice say words with differing emotional intonations. The words were all foods, with the voice either expression like or dislike for the food through the intonation. The children heard the voice then clicked on either the happy or sad face (seen above) as to how they interpreted the emotion of the voice.

B.3.7. Expressive Affect

Figure B.13 PEPS-C 2015 Expressive Affect - Stimulus: Children were first presented with an image of a food item. They were asked the name the food item, but to express with their voice whether they liked or disliked the food (like or dislike intonation, as in the Receptive Affect task).

Figure B.14 PEPS-C 2015 Expressive Affect - Response: Following the child saying the food item, they were scored by the researcher using the keypad, indicating whether they felt the child liked the food (press ☺), disliked the food (press ☹), or the response was ambiguous (A). Finally the child clicked one of the faces as seen above to confirm in the system whether they said they did or did not enjoy the food item. The children scored correctly if the researcher's response matched their own.

B.3.8. Receptive Lexical Stress

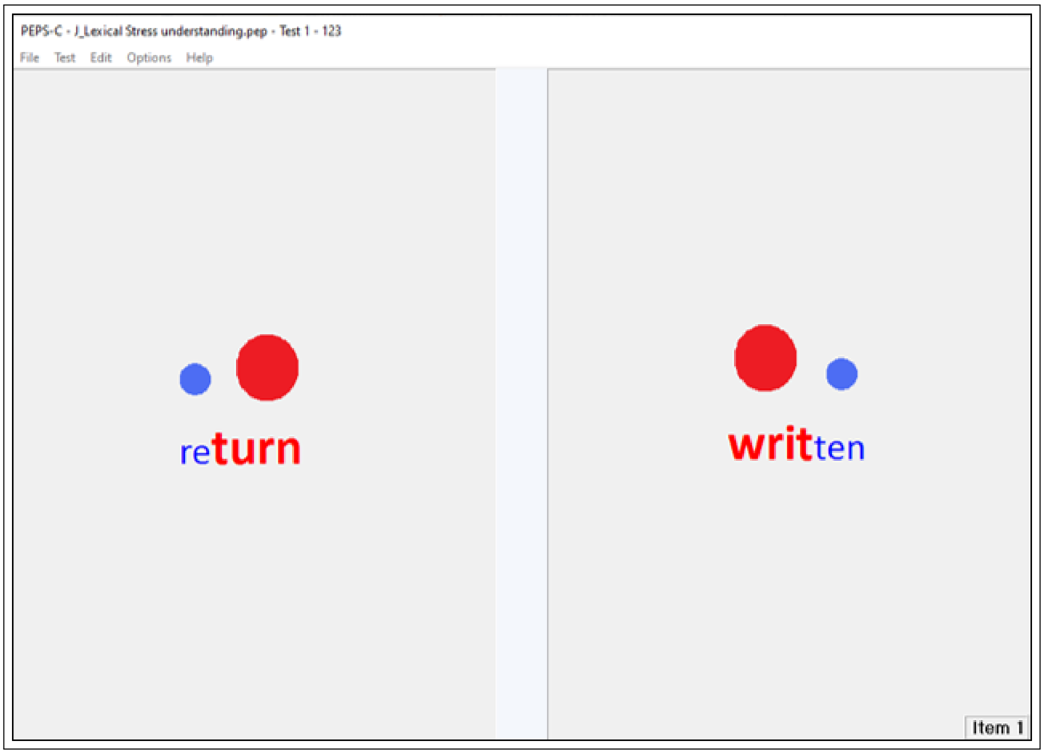

Figure B.15 PEPS-C 2015 Receptive Lexical Stress Task: Children heard the voice on the computer say a word. The stimuli were matched so that, depending on where the stress was placed, it would have a different meaning. Children listened carefully and then clicked on one of the two words on the screen which they felt represented what they heard (i.e. where the main stress was placed).

B.3.9. Expressive Lexical Stress

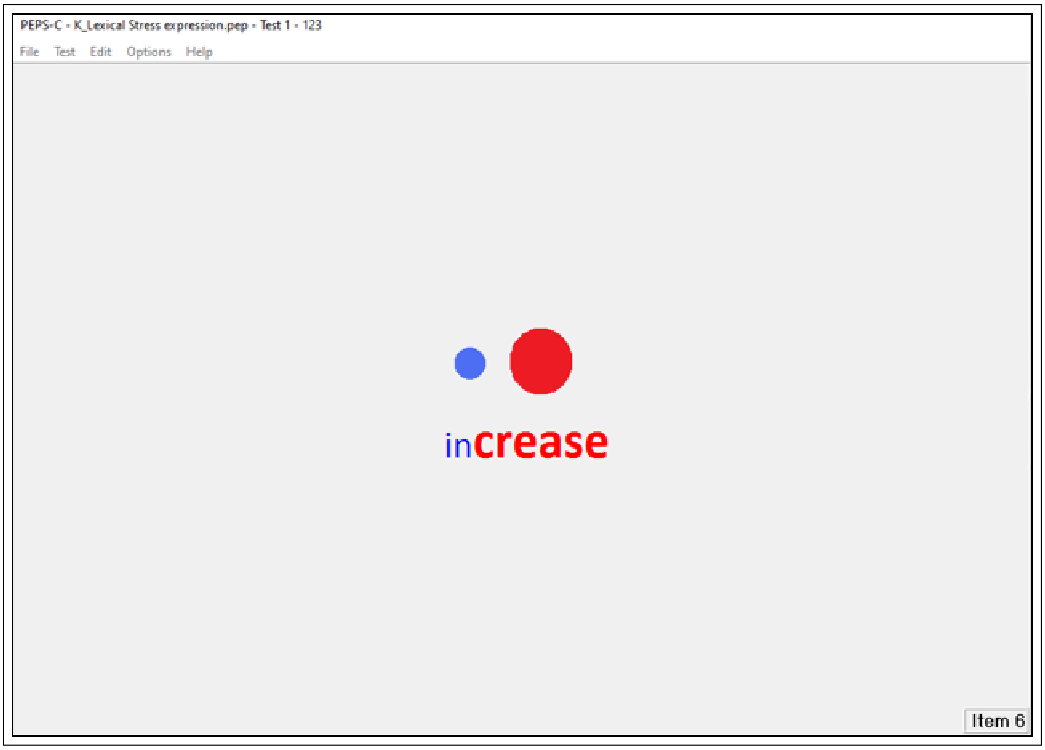

Figure B.16 PEPS-C 2015 Expressive Lexical Stress Task: For this task the children were presented with a word onscreen. The word was highlighted to emphasise the syllable to be stressed. . The child then produced the word with the appropriate stress, and the researcher recorded whether they thought the child's stress was on the first syllable (press 1), second syllable (press 2) or an ambiguous response (press A).

B.3.10. Receptive Phrase Stress

Figure B.17 PEPS-C 2015 Receptive Phrase Stress Task: This task was to investigate the ability of children to comprehend differences in phrase stress. Children heard a pre-recorded voice speak a short phrase, using either the single word (e.g. wetsuit) or two-word (e.g. wet suit) version, which required understanding of the differing stress patterns for each. Once they heard the phrases, children clicked whether they heard the two-word phrase or compound noun within the phrase they just heard.

B.3.11. Expressive Phrase Stress

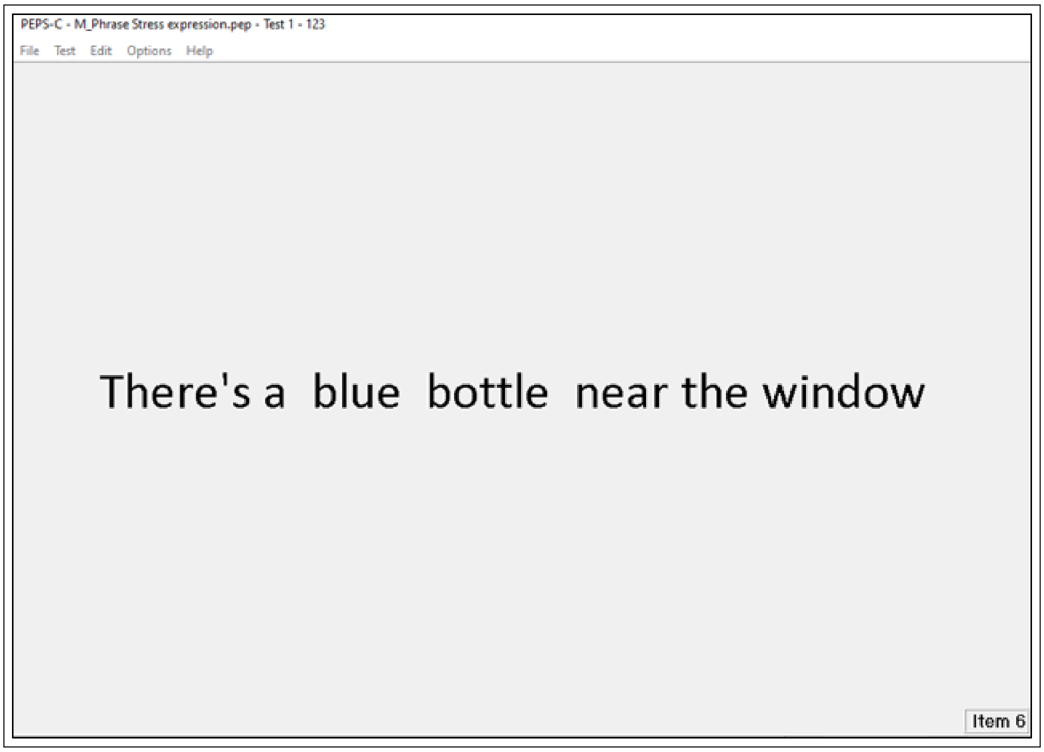

Figure B.18 PEPS-C 2015 Expressive Phrase Stress Task: Children were presented with short phrases on the screen. The child read the sentences shown on screen, and the researcher recorded whether they thought the child used a compound noun (press 1), a two-word phrase (press 2) or it was ambiguous (press A).

B.3.12. Receptive Boundary

Figure B.19 PEPS-C 2015 Receptive Boundary Task: This task looked at how well children could disambiguate phrases that were syntactically ambiguous. The pre-recorded sentences were spoken with a prosodic boundary after the first word (e.g. 'pink / and black-and-green socks') or after the second word (e.g. 'pink-and-black / and green socks'). Children listened to these phrases and selected the picture which best represented what they heard.

B.3.13. Expressive Boundary

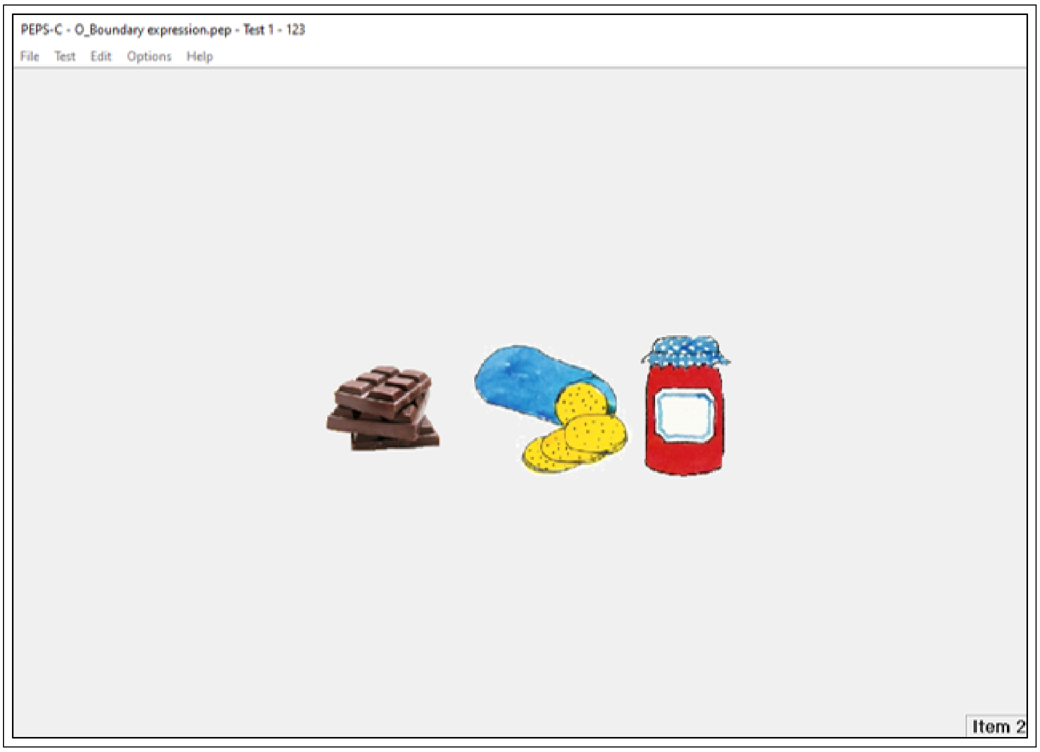

Figure B.20 PEPS-C 2015 Expressive Boundary Task: This task investigated children's ability to understand and use appropriate prosodic boundaries. For example, this picture shows pictures of chocolate, biscuits and jam. The appropriate response would be "chocolate, biscuits and jam", not "chocolate biscuits and jam". The researcher recorded on the keypad whether children marked the boundary after the first word (press 1), after the second word (press 2), or if the boundary was ambiguous (press A).

B.3.14. Receptive Contrastive Stress

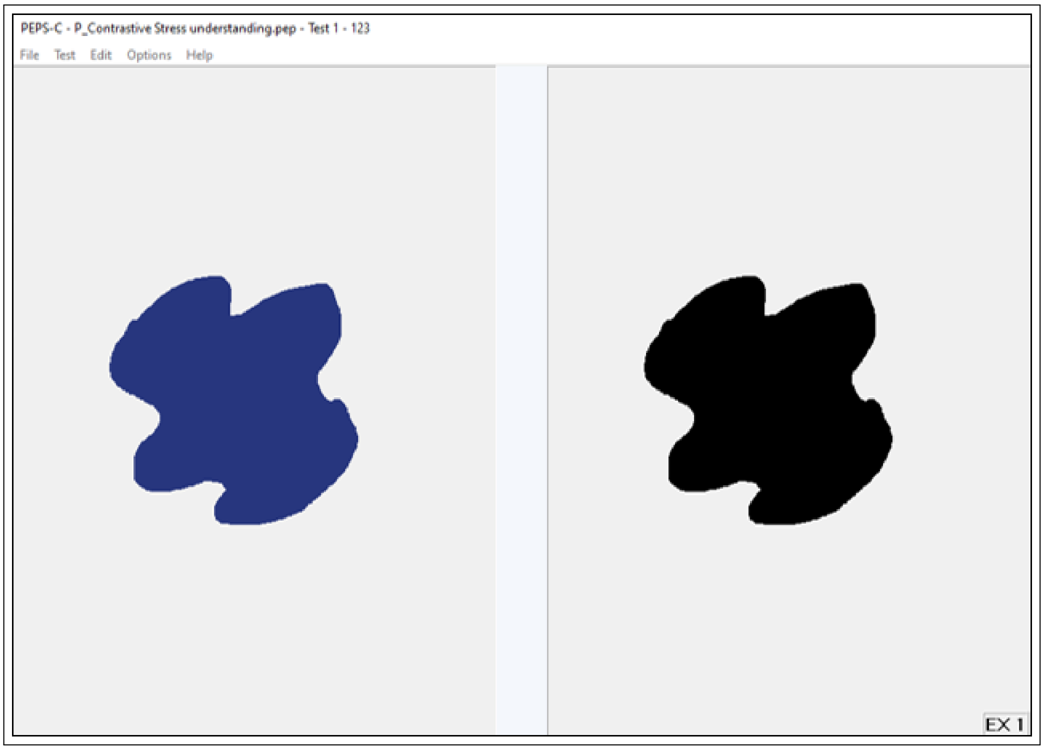

Figure B.21 PEPS-C 2015 Receptive Contrastive Stress Task: This task looked at children's ability to understand contrasting stress when used in short sentences. The conceit was that the speaker had bought some socks but had forgotten to buy one colour, with the stressed word indicating the colour forgotten (e.g. "I wanted blue and BLACK socks"). Children had to listen to these short sentences, and select from two colours shown on the screen as to which colour socks they thought the speaker had forgotten to buy.

B.3.15. Expressive Contrastive Stress

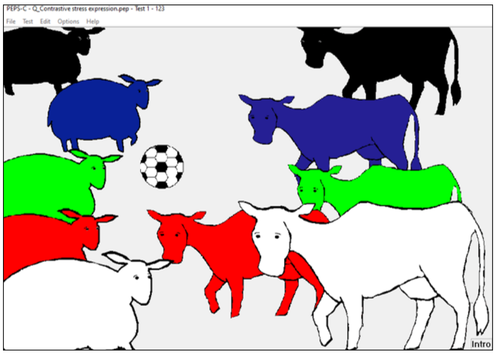

Figure B.22 PEPS-C 2015 Expressive Contrastive Stress - Setup: This slide allowed the researcher to explain the concept of the two animal teams playing football, and point out the colours of the animals, ready for the actual task.

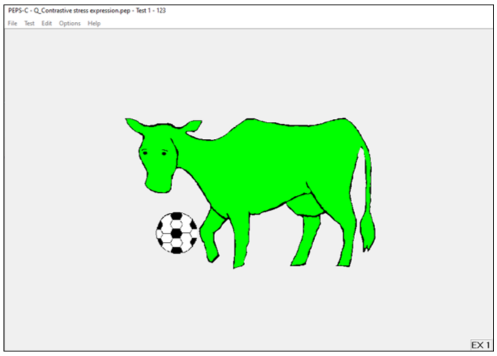

Figure B.23 PEPS-C 2015 Expressive Contrastive Stress - Stimulus: This final task was the expressive variant of the previous task, and required children to listen to the recorded voice narrate a football match between sheep and cows of various colours (black, blue, green, red and white). The recorded voice narrated the match, and a picture on screen represented what the speaker was commenting upon. However, the narrator would get one aspect of the commentary wrong each time, either the animal with the ball or its colour. The child was then required to correct what had been said, using appropriate contrastive stress to ensure the researcher understand the difference. As an example, there is a picture of a green cow with a football on screen, the narrator says "The green sheep has the ball", then the appropriate response would be "No, the green COW has the ball". The researcher listens to the child say these corrective sentences and marks where they believe the child placed the contrastive stress, either on the colour (press 1), the animal (press 2) or if the stress was ambiguous (press A).

B.4. Test of Word Reading Efficiency - Second Edition (TOWRE-2)

B.4.1. TOWRE-2 Short Word Reading

Practice Words	then	natural
on	spell	purchase
my	come	vacant
bee	start	everyone
old	green	swollen
warm	want	fireplace
bone	better	together
most	learn	horizon
spell	black	embassy
	train	mountain
Test Words	even	project
go	went	factories
dog	thing	straighten
in	other	clarify
at	fruit	frequent
am	wrong	mediate
it	watch	threshold
so	truck	modulate
big	stars	prudent
be	winter	exercise
do	begin	protect
box	forest	desperate
one	street	quantity
look	chance	wonderful
if	instead	initiate
not	farmer	spurious
car	spring	particular
hot	present	emergency
this	peace	selection
have	huge	verbatim
some	believe	awkward
now	office	wilderness
need	question	grandiose
give	contact	ornament
sat	history	penitent
good	invent	component
here	invoice	heritage
shop	complete	sceptical
meat	custom	transfusion
best	inquire	

Figure B.24 TOWRE-2 Short Word Reading Word List: List of words children were asked to read during the 45 second time limit.

B.4.2. TOWRE-2 Phonemic Decoding Efficiency

Practice Words
ba
um
fos
gan
rup
nasp
luddy
dord

Test Words
Mo
Ik
Pu
Bi
Ib
Ku
Eb
Pog
Dat
Mip
Ral
Nas
Mib
Faw
Shum
Bice
Nade
Teap
Derl
Marl
Berk
Mest
Stree
Weaf
Barch
Glack
Prot
Runk
loast
mact
blork
phet
wogger
klup
skad
keast
churl
glamp
prait
flact
throbe
creft
flimp
girtus
strale
debmer
happon
framble
progus
supken
jeltlic
tegwop
slinperk
plinders
thundelp
bramtich
chimdruff
darlankert
stremflick
morlingdon
revignuf
obsorfelm
pitocrant
glimpobot
strilmolifant
bormorint

Figure B.25 TOWRE-2 Phonemic Decoding Efficiency Word List: List of nonwords children were asked to reading during the 45 second time limit.

B.5. Test for Reception of Grammar (TROG-2)

B.5.1. TROG-2 Modified - Listening Comprehension

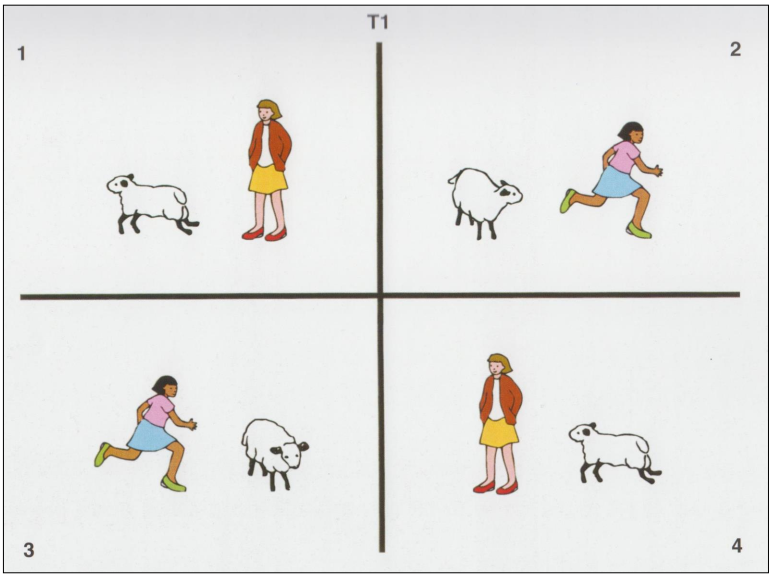

Figure B.26 TROG-2 Modified - Listening Comprehension Example Stimulus: This figure is an example of one of the TROG-2 Listening Comprehension stimuli, which was displayed to children on printed pieces of A4 paper. The researcher then read a sentence out (e.g. "The sheep the girl looks at is running"), and the child pointed to the one picture they thought the sentence referred to. There were 42 pictures shown in total, all of the sentences that could be noun-swapped with a matching picture were included.

B.5.2. TROG-2 Modified - Reading Comprehension

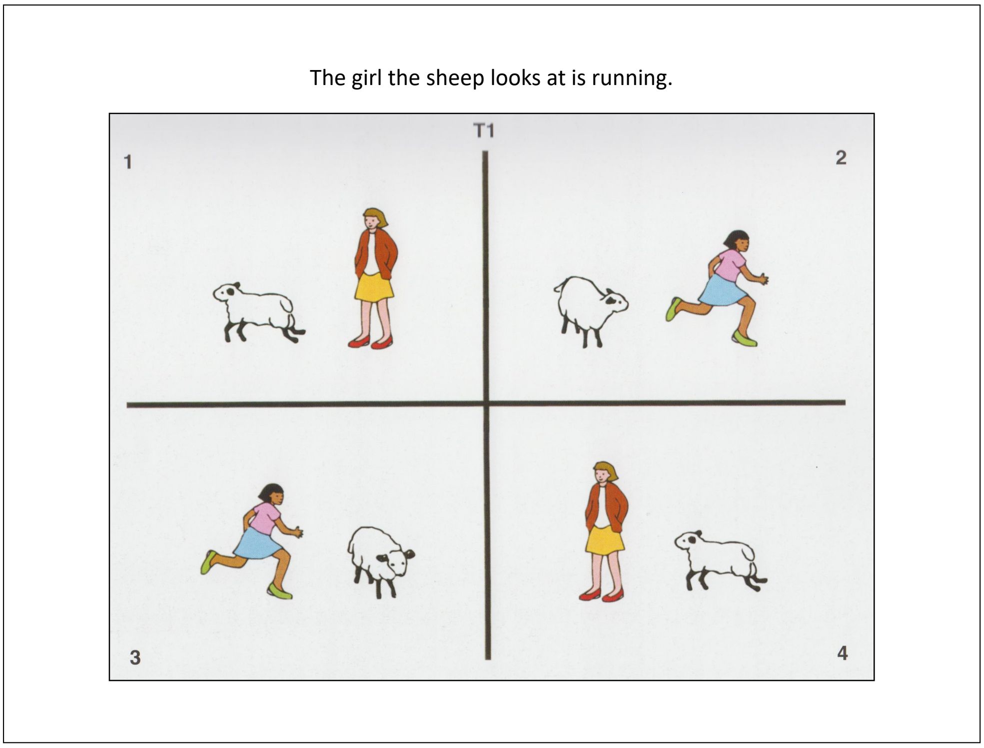

Figure B.27 TROG-2 Modified - Reading Comprehension Example Stimulus: This figure is an example of the modified version of the TROG-2 stimuli. The same 42 sentences were used as in the listening comprehension version, but this time they were noun-swapped. Furthermore, the sentence to be read was written at the top of the page, which the child was asked to read out loud, and then to point to the picture they thought the sentence was referring to.

B.6. York Assessment of Reading for Comprehension (YARC)

B.6.1. YARC Short Word Reading Test

Block 1
see
look
play
was
like
this
next
house
going
bell

Block 2
hang
stand
their
living
again
first
slowly
score
found
bread

Block 3
scream
journey
suppose
yawned
should
tissue
caught
stretching
tongue
copies

Block 4
medicine
strengthen
source
creative
material
eventually
hygiene
despite
calm
journalism

Block 5
excitable
dehydration
persuade
aggrieved
originate
courageous
atmospheric
familiarise
scenic
recurrence

Block 6
ferocious
cynical
excursion
coincidental
abysmal
endeavour
rheumatism
haemorrhage
liaise
pseudonym

Figure B.28 YARC Short Word Reading Test Word List: List of words children were asked to read. These were split into blocks of ten, and get progressively more difficult as the child moves down the page.

B.6.2. YARC Reading Comprehension: Passage Reading Primary (age 5-11)

The house was dark and no one was home. A burglar was in the bushes. He crept around to the back of the house and peered through the window. Carefully he broke the glass.

Once inside, he gazed at the contents of the room. He slipped some silver from the mantelpiece under his overcoat and found some jewellery in a drawer. Suddenly, he heard a key in the door. The burglar quickly squeezed under the large desk. A man walked into the house. He put his briefcase down and went upstairs.

The burglar decided to get out quickly. As he crept towards the door he tripped over something. The dog was startled and began barking loudly. A light came on upstairs. Before the burglar could escape, he felt the man's steely grip on his shoulder. The man had a truncheon and handcuffs. Unfortunately for the burglar, he had broken into a policeman's house.

A–3

Figure B.29 YARC Reading Comprehension: Passage Reading Primary (age 5-11) - Fiction Passage: Taken from the the YARC Passage Reading Booklet, this text was the fiction text chosen from Form A of the booklet.

In Australia, the reptiles that are known worldwide as monitor lizards, are called goannas. Goannas range in length from 20 centimetres to over 2 metres, but all have the same distinctive shape: a flattened body, long neck with loose skin under the throat, strong legs with long toes and sharp claws, and a long tail.

Mostly they are ground dwellers, hunting close to the burrows in which they live, but they are also good tree climbers and strong swimmers. Being carnivores, they eat lizards, snakes, small mammals, birds, and eggs; often swallowing the animal whole. They hunt their prey by tracking and attacking it, or by using their sharp claws to excavate animals and eggs hidden in the ground.

Like most lizards, goannas lay eggs, usually in a nest or burrow and between seven to thirty-five eggs at a time. These hatch in eight to ten weeks.

Although bulky in appearance, goannas can run swiftly on two legs. They also rear up on two legs when threatened, inflating flaps of skin around their throats, hissing, and lashing out with their powerful tails.

A – 4

Figure B.30 YARC Reading Comprehension: Passage Reading Primary (age 5-11) - Non-Fiction Passage: Taken from the YARC Passage Reading Booklet, this text was the non-fiction text chosen from Form A of the booklet.

Level 3: Burglar

Question	Correct responses Score 1	Incorrect responses Score 0
1. Outside, where did the burglar hide?	In the bushes	Q Outside the house In the back garden Under a big desk
2. How did the burglar get into the house?	He broke the glass and got in through the window Because he broke the window	Q He broke the glass Q Through the window
3. What did he try to steal?	Silver **and** jewellery	Q Silver Q Jewellery Money The mantelpiece
4. When did the burglar hide under the desk?	When he heard the man unlocking the door When he heard the key in the door	Q When the man came in Q When he heard the man
5. Who went upstairs?	The owner The policeman The man that lived in the house	Q The man The burglar
6. Why did the owner come downstairs?	Because the dog was barking He wanted to know why the dog was barking	Q To see what was happening Q To catch the burglar Q He heard a noise
7. What does 'steely grip' mean?	A really hard grip He held onto his shoulder very tight Really holding it hard so he couldn't get away	Q He's trying to hold the man still with his grip They grab you It feels like steel
8. When do you think this story takes place?	Night time Evening	Q When the burglar breaks in Outside In the policeman's house In the afternoon

Figure B.31 YARC Reading Comprehension: Passage Reading Primary (age 5-11) - Fiction Passage Questions and Answers: Taken from the the YARC Passage Reading Manual, this was the list of questions for passage A-3, with the answers and incorrect responses with possible clarification criteria (incorrect responses marked with a Q allowed the researcher to use a pre-planned prompt for the children to see if they have a correct answer).

Level 4: Lizards

Question	Correct responses Score 1	Incorrect responses Score 0
1. What is a goanna commonly known as?	Monitor lizard	Q A lizard A reptile Worldwide lizard
2. What sort of homes do goannas live in?	Burrows	Q Underground holes
3. Why would it be wrong to say that goannas only stay on or under ground?	They climb trees They are good swimmers	They stay underground They run fast
4. How do you know that goannas are carnivores?* **Do not explain what a carnivore is.*	Because they eat any of: lizards, snakes, small mammals, birds, eggs	Q It says so Q They can eat animals Q They are meat eaters They are reptiles
5. Name one thing for which the goanna would use its long toes and claws.	Excavate/dig up small animals and eggs from the ground To attack with For climbing trees	Q For digging For tracking For running
6. How long do goanna eggs take to hatch?	8 to 10 weeks	Q Up to 10 weeks 7 to 10 weeks 8 to 10 days
7. When would a goanna hiss and inflate the skin around its throat?	When it is threatened When something is trying to attack it	When they catch their prey When it is attacking
8. How do you know that goannas are not slow, clumsy creatures?	They can run swiftly on two legs They can run fast	They rear up on two legs They attack quick

Figure B.32 YARC Reading Comprehension: Passage Reading Primary (age 5-11) - Non-Fiction Passage Question and Answers: Taken from the YARC Passage Reading Manual, these were the list of questions for passage A-4. As above, correct answers and incorrect responses were detailed, with those incorrect responses marked with a Q present for allowing children a chance to answer correctly.

Appendix C. Results Details

C.1. R Scripts

We provide a hyperlink to the R scripts used to analyse the data for this project. The scripts were written and processed in R Studio with R Markdown scripts, for ease of output and reading. We have prevented packages from installing when using the scripts, allowing the user to choose whether they wish to install these packages with their own consent. Also included is the .csv file with the anonymised raw data for each child, this was the file input to R for cleaning and analysis.

https://www.dropbox.com/sh/e8z01odhjmgy5ul/AAAYPcnIti13TkyfQM14aSVqa?dl=0

C.2. Visualisations of Normality

C.2.1. CTOPP-2

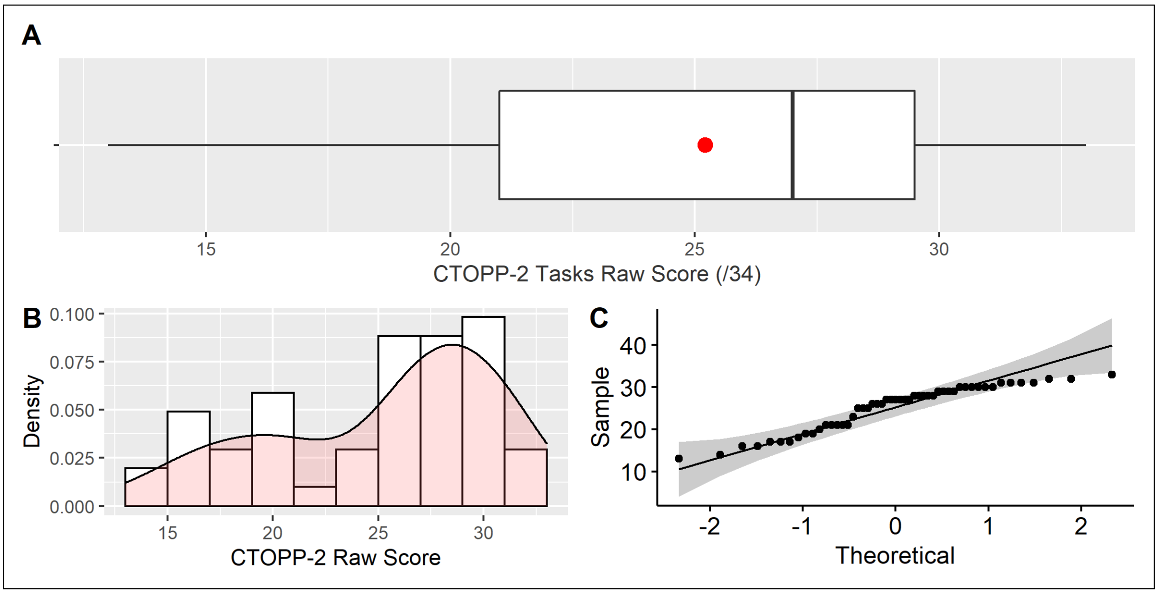

Figure C.1 CTOPP-2 Elision Visualisations: A = Boxplot, B = Histogram and Density Plot, C = Q-Q Plot.

C.2.2. Digit Span

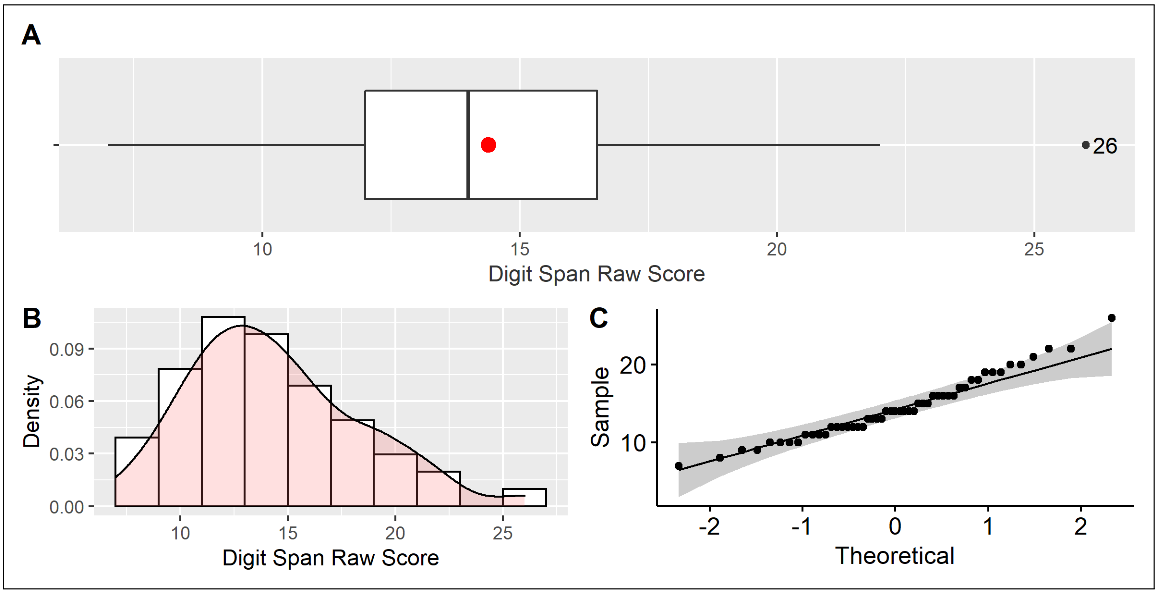

Figure C.2 Digit Span Visualisations: A = Boxplot, B = Histogram and Density Plot, C = Q-Q Plot.

C.2.3. PEPS-C 2015: Expressive Tasks

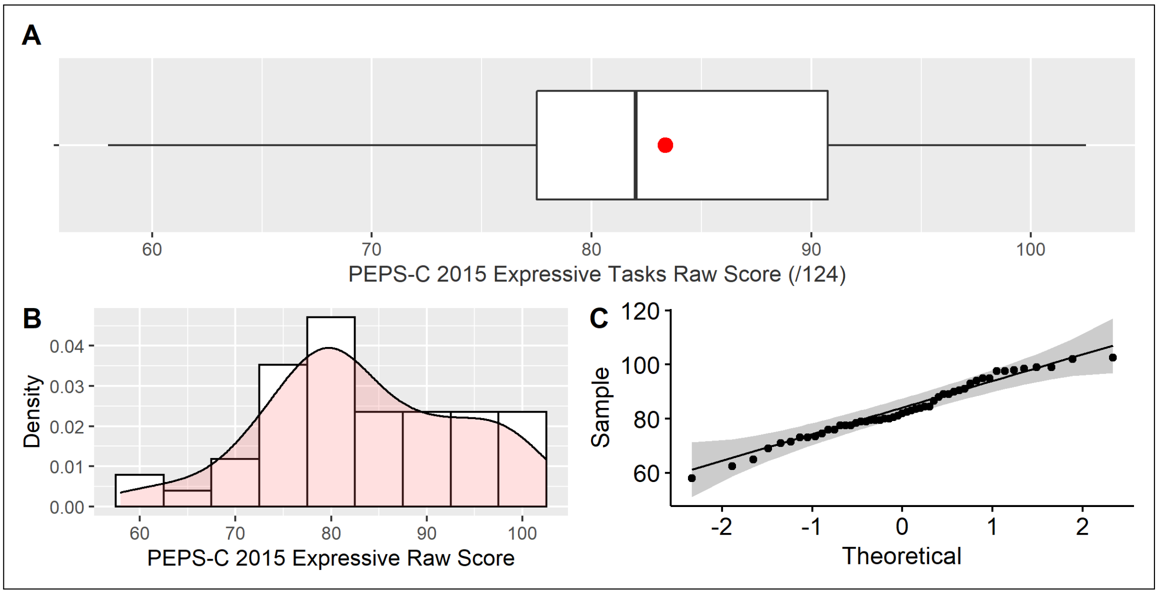

Figure C.3 PEPS-C 2015: Expressive Visualisations: A = Boxplot, B = Histogram and Density Plot, C = Q-Q Plot.

C.2.4. PEPS-C 2015: Receptive Tasks

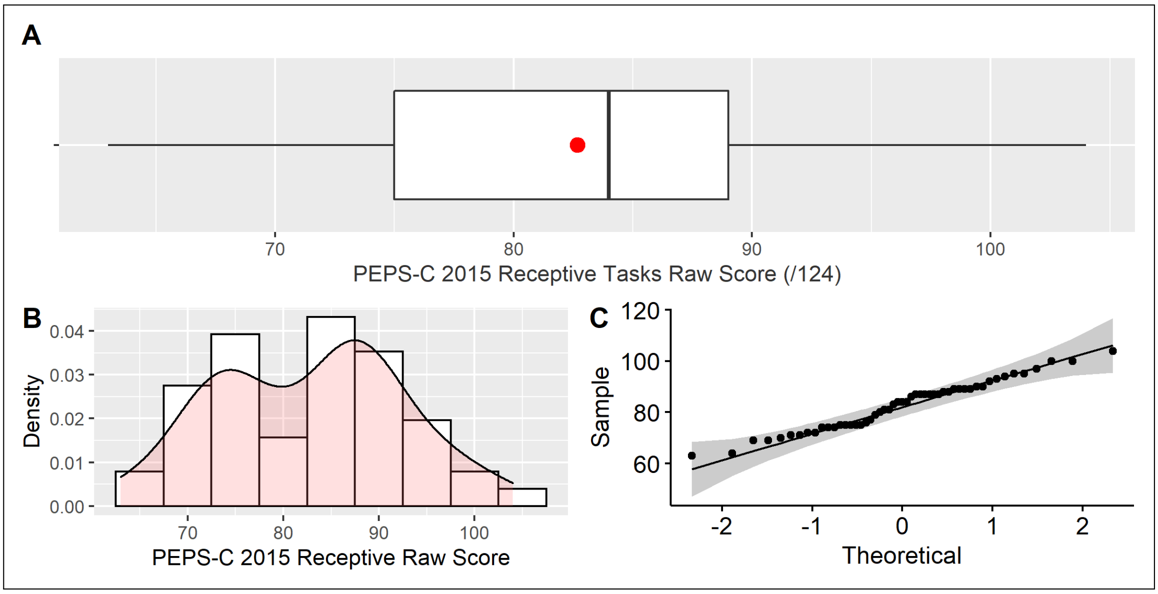

Figure C.4 PEPS-C 2015: Receptive Visualisations: A = Boxplot, B = Histogram and Density Plot, C = Q-Q Plot.

C.2.5. TOWRE-2 PDE

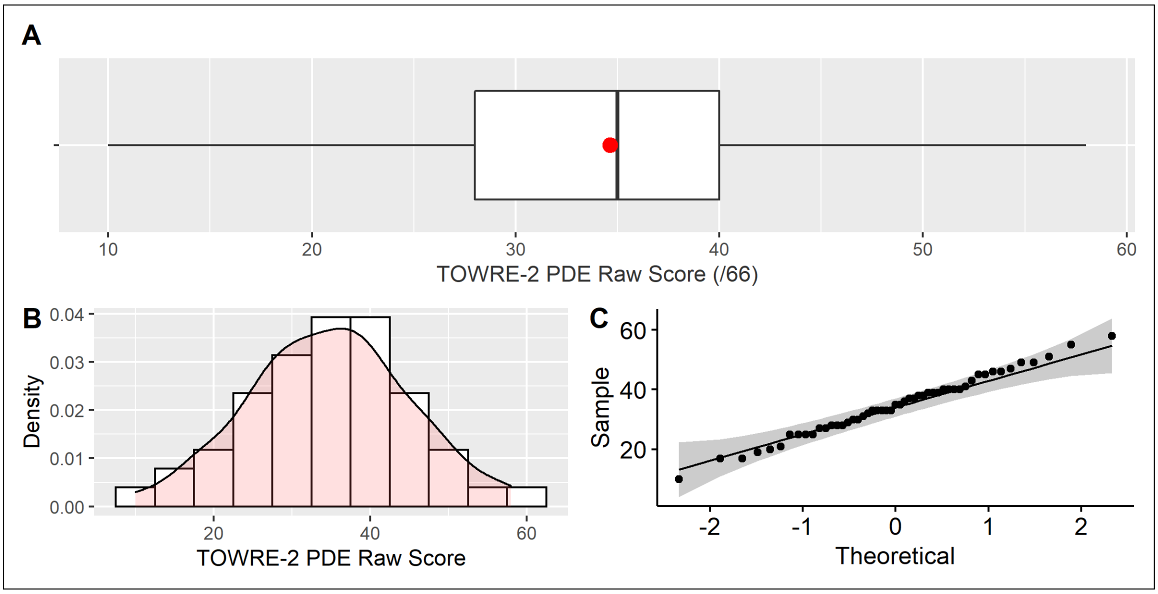

Figure C.5 TOWRE-2 PDE Visualisations: A = Boxplot, B = Histogram and Density Plot, C = Q-Q Plot.

C.2.6. TOWRE-2 SWR

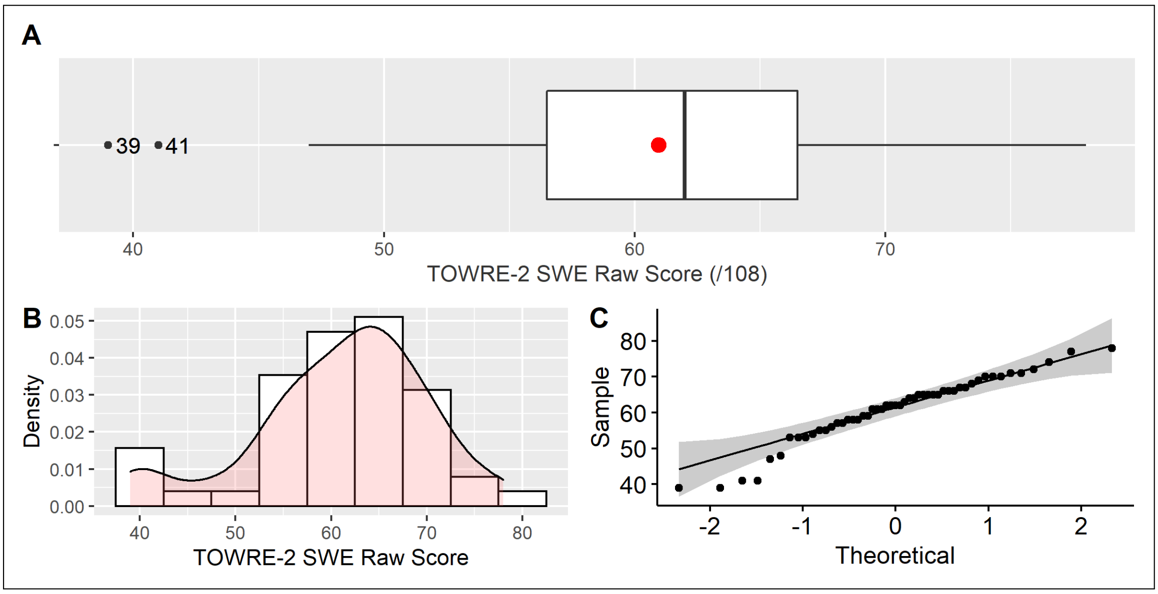

Figure C.6 TOWRE-2 SWE Visualisations: A = Boxplot, B = Histogram and Density Plot, C = Q-Q Plot.

C.2.7. TROG-2 LC

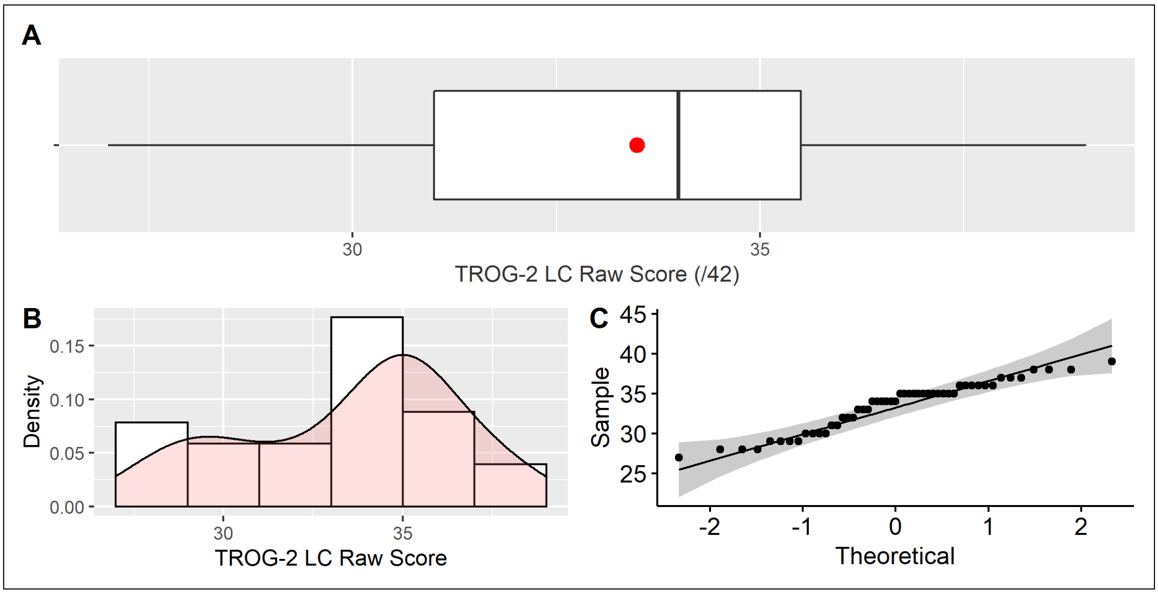

Figure C.7 TROG-2 LC Visualisations: A = Boxplot, B = Histogram and Density Plot, C = Q-Q Plot.

C.2.8. TROG-2 RC (Modified)

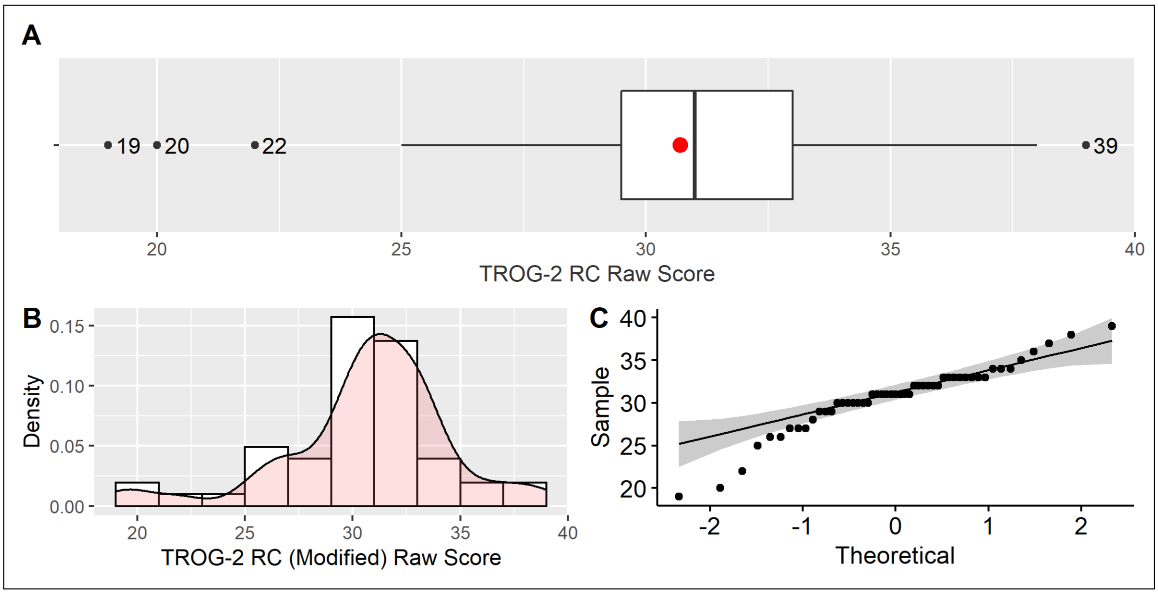

Figure C.8 TROG-2 RC Visualisations: A = Boxplot, B = Histogram and Density Plot, C = Q-Q Plot.

C.2.9. YARC SWR

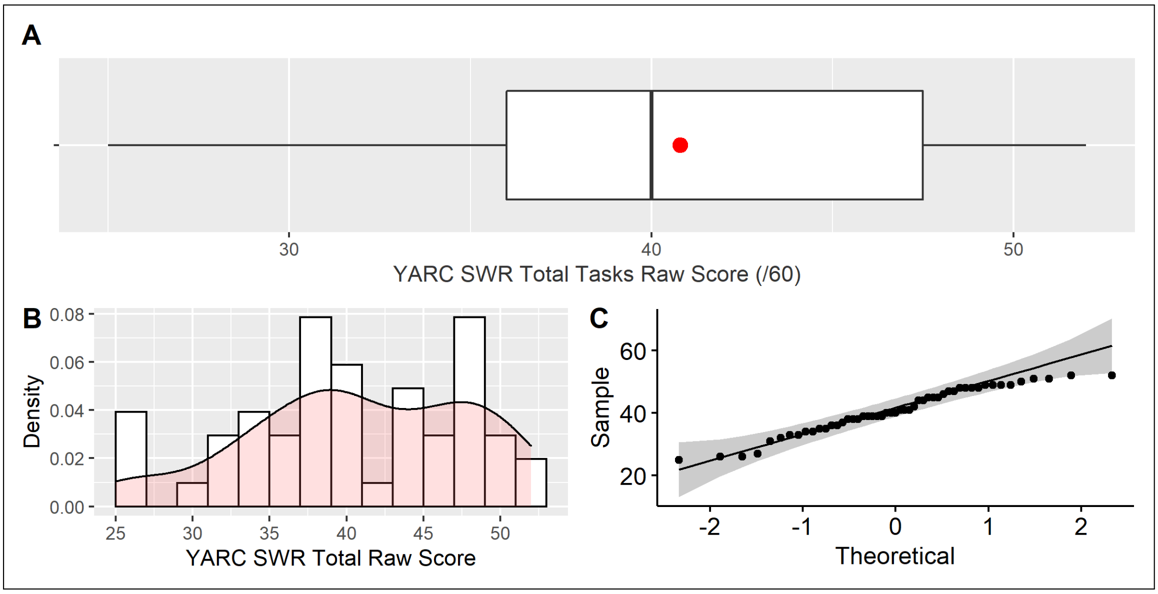

Figure C.9 YARC SWR Visualisations: A = Boxplot, B = Histogram and Density Plot, C = Q-Q Plot.

C.2.10. YARC Passage Reading

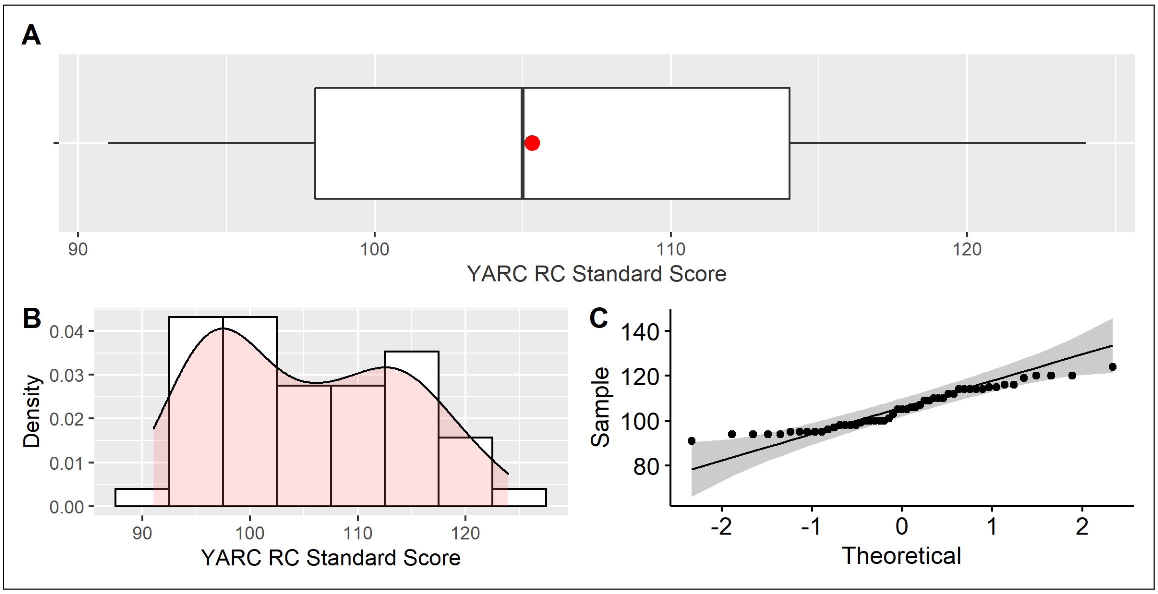

Figure C.10 YARC RC Standard Score Visualisations: A = Boxplot, B = Histogram and Density Plot, C = Q-Q Plot.

C.3. Visualisations of Transformed Variables

C.3.1. Transformation of TROG-2 Reading Comprehension (Modified)

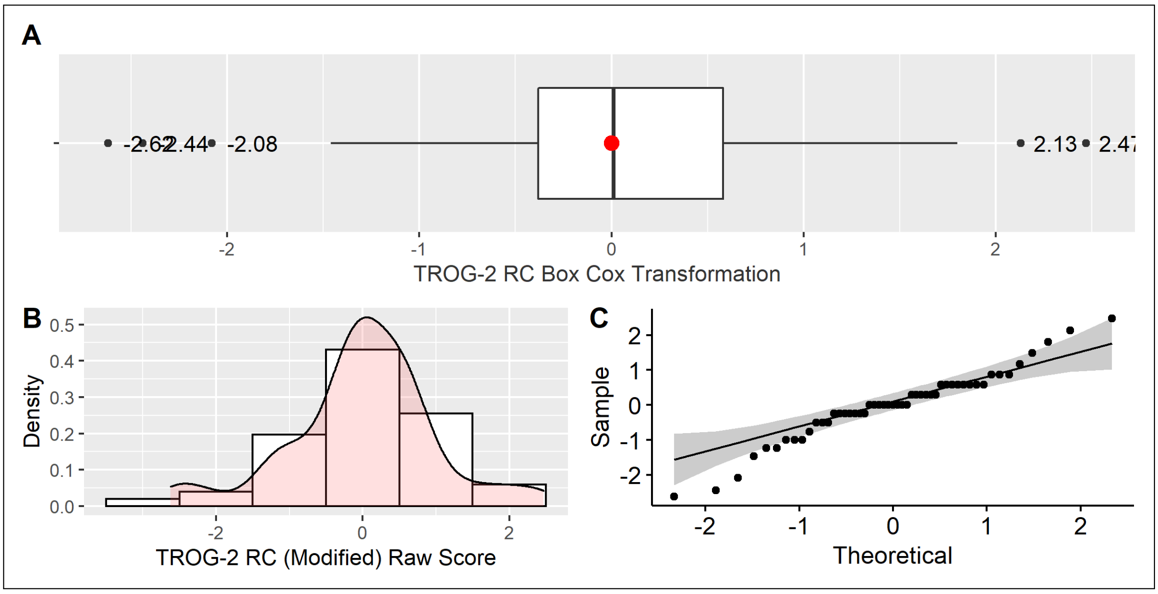

Figure C.11 TROG-2 Reading Comprehension (Modified) Tranformed Visualisations: A = Boxplot, B = Histogram and Density Plot, C = Q-Q Plot.

C.4. Data Cleaning and Assumptions

In this section, extra details regarding data cleaning and the assumption tests are listed. For the main body of the thesis it was felt these details would be too much to read, so the discussion has been moved to hear if clarification for these decisions is required in more detail.

C.4.1. Outliers

In choosing whether to delete or retain outliers, we used the differences in the mean as an aid. The mean as a measure of central tendency is the most affected by extreme scores, and as such it helps to look at the difference in means between the datasets with and without the outliers. If we feel the difference in mean is too extreme, then a major consideration regarding deletion of outliers or transformation of the dataset should be had.

Taking the Backwards Digit Span task first, there was only one identified outlier in the dataset, and looking at the difference in the means there is no major change. As such, we decided to not take out the outlier from here. The outlier was a child who scored the highest on the digit span task, they simply got one step further than the other children to memorising six numbers and recalling them backwards. As such, this was not a mistake, simply that one child was particularly good with their short term working memory to complete this task.

Next, is the TROG-2 RC (Modified) outliers. Despite having a greater number of outliers, the mean difference between their inclusion and exclusion does not change radically, especially if you round to a whole number where you find they are the same. The influence of the outliers can be seen more clearly when looking at the distribution of the datapoints when we visualise the data through a histogram and density plot (see Appendix C.2), but taken from just these values there seems no immediate need to delete the outliers. The reason for these outliers is likely due to the difficulty some found with the task, which involved reading sentences with complicated grammatical constructions.

Lastly, we have the TOWRE-2 SWE variable, the real word reading task where children read as many real words as they could in 45 seconds. The analysis highlighted four outlying data points here, and their removal increased the mean from 61 to 63, a two point increase in the mean value. We compared the correlation analyses with and without the outliers and found no significant difference in the relationships, so decided to keep the task the same.

The decision has been made not to remove the outliers at this stage, as from closer analysis these are not measurement errors on part of the researcher who collected the data, nor are they data entry mistakes. These extreme values are due to individual differences in ability it seems, rather than an erroneous measurement. To account for this, we ran each measurement with and without the outlier to see if there was any major change in the output. The expectation was a small, but insignificant change in the results. Furthermore, as discussed in ref there may be more significant transformation of the data related to issues of normality, skewness and kurtosis. Through these diagnoses and alterations the issue of the outliers may be altered in step with these decisions.

C.4.2. Transformation of Data

The decision to transform a variable is one which requires serious consideration. The utility of transformation is helpful for correcting our issues with normality for conducting our data analysis. Regarding our own dataset, we are not comparing changes in measurements over time, so we are able to change a single variable when looking at relationships between variables, rather than transforming every measured variable. Though not changing the fundamentals of the variables (for example if we use the square root transformation, we can then square the output to return our variables to their original form) transformation is not agreed to be a good idea. Different schools of thought feel different about transformation (a discussion of which is not fully explored in this thesis), and the use of robust statistics and transformations such as Tukey's Ladder of Power and the Box-Cox transformation have been looked at. Where we have decided to apply transformations to the data will be noted through the rest of this chapter, we computed the outputs for these measures with both the transformed and non-transformed variables to see what difference this made.

To aid the choices of transformation we utilised an R package suitably named *bestNormalize* created by Peterson and Cavanaugh (2019). This package was created to centralise a number of transformations for data that attempt to transform data towards the Gaussian distribution of data, the normal distribution. An advantage of this package is that the output created includes both the pre-transformed and transformed variables in a list for reference, allowing for an easy way to reverse the data between the two sets and report the output with the original values. The function we used from the package is coincidentally also called *bestNormalize*, and as its name suggests attempts to select and inform you of the best transformation to use for your data. This is done via the use of the Pearson P statistic (divided by the degrees of freedom), and the use of the output is a relatively good goodness of fit test between the different transformations. The test ratio closest to 1 is the transformation which follows closest to a normal distribution. An important aspect of this package is that the functions are running randomly generated repetitions, so essentially re-running the function *bestNormalize* could change the recommended transformation chosen. However, the change in values was usually so small to not make much difference, this is likely because of the relatively small sample size we have for the current project, and so the transformations we chose were of personal preference for the data. Furthermore, the authors recognise this fact with the small sample sizes, but add that the transformations with small sample sizes should look similar in size to each other so should not be a major deal breaker in this instance. To try and control this as much as possible, we used the *set.seed* function in R to control the variation as much as possible. When we identified the normalization transformation we wanted to use, we commented out the *bestNormalize* function, as re-running the script would change the choice, and re-ran the transformation with the specific transformations function.

The most major violation of the the normality assumptions, according to the Shapiro-Wilk statistic, was the CTOPP-2 Elision measure. Visualising the measure this distinction could be seen by the negative skew and the dip in scores in the middle of the histogram. As such, a transformation was used to help transform the data closer to the gaussian distribution.

The Box-Cox transformation was used for the TROG-2 Listening Comprehension variable, as it showed a slight non-normal distribution identified by the Shapiro-Wilk statistic. We decided no transformation was appropriate for this point, as when using the *bestNormalize* function the resulting Pearson P outcome indicated that no transform was just as effective as the other transformations, suggesting transformation would only bring the smallest of benefits to the dataset. This was confirmed when we used the Box-Cox transformation and found an improvement in skewness, but not in kurtosis, and an insignificant change in shapiro-wilk value. At this stage in the data analysis we did not transform the data, we considered it near enough to normal to be happy using it for the Pearson correlation test, and would see what happened with the regression when we checked the assumptions.

We also used the Box-Cox transformation for the TROG-2 RC (Modified) variable, which selected $\lambda = 2$ as the appropriate power transformation for the dataset and then centered and standardised the variable for legibility. This makes sense due to the significantly negative skew and kurtosis we identified. The expectation was not to get a perfect gaussian distribution, but make an attempt for normality due to the range of scores in this task. The visualisation of this transformation can be seen in Figure (C.11), and it can be seen on the histogram that the negative skewness has been lessened from where it was prior (supported by a new skewness statistic of -0.261) and kurtosis below 1 at 0.757. It is important to reiterate that these statistics are not foolproof, but along with the visualisation the distribution appears closer to normal and more appropriate for the parametric statistics we used. The imperfection can be seen via the Q-Q plot, where we see that the lower and upper tails do deviate from the expected line, and the outliers in the boxplot more than 2 standard deviations from the mean when standardised. In the end, we did not use this transformed variable, we used other methods of transformation (e.g. Weighted-Least Squares modelling) to account for the TROG-2 RC issues.

Finally, the YARC Reading Comprehension (Standard Score) variable. We used the *bestNormalize* function once more to aid us in deciding upon the transformation to conduct. To our surprise, it was suggested that conducting no transformation would be just as good as performing any of the other choices we had. This may be due to the skewness of this dataset not being too high, only that the kurtosis was short-tailed. However, when visualising the plots we do see the bumps in the density curve, but the spread of the data is likely due to being a standard score, wherein children's scores are transformed related to their age. We would expect this high variations of ability on reading. As such, we decided not to transform the data initially, and will check the assumptions of the regressions and mediation later on to ensure this does not cause an error.

References

Ahmed, Y., Francis, D. J., York, M., Fletcher, J. M., Barnes, M., and Kulesz, P. (2016). Validation of the direct and inferential mediation (dime) model of reading comprehension in grades 7 through 12. *Contemporary Educational Psychology*, 44–45:68–82.

Anderson, R. (1994). Role of the reader's schema in comprehension, learning, and memory. In Ruddell, R., T., R., and Singer, H., editors, *Theoretical Models and Processes of Reading*. International Reading Association, Newark.

Aouad, J. and Savage, R. (2009). The component structure of preliteracy skills. *Canadian Journal of School Psychology*, 24(2):183–200.

Arciuli, J. (2017). The relationship between children's sensitivity to dominant and non-dominant patterns of lexical stress and reading accuracy. *Journal of Experimental Child Psychology*, 157:1–13.

Arciuli, J. and Ballard, K. J. (2017). Still not adult-like: Lexical stress contrastivity in word productions of eight- to eleven-year-olds. *Journal of Child Language*, 44(5):1274–1288.

Ashby, J. (2016). Why does prosody accompany fluency? Re-conceptualizing the role of phonology in reading. In Khateb, A. and Bar-Kochva, I., editors, *Reading Fluency*, pages 65–89. Springer International Publishing, Switzerland.

Ashby, J. and Clifton, Jr., C. (2005). The prosodic property of lexical stress affects eye movements during silent reading. *Cognition*, 96(3):89–100.

Ashby, J., Dix, H., Bontrager, M., Dey, R., and Archer, A. (2013). Phonemic awareness contributes to text reading fluency: Evidence from eye movements. *School Psychology Review*, 42(2):157–170.

Bader, M. (1998). Prosodic influences on reading syntactically ambiguous sentences. In Fodor, J. and Ferreira, F., editors, *Reanalysis in Sentence Processing*, pages 1–46. Springer, Dordrecht.

Baron, R. M. and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *Journal of Personality and Social Psychology*, 51(6):1173–1182.

Beattie, R. L. and Manis, F. R. (2013). Rise time perception in children with reading and language difficulties. *Journal of Learning Disabilities*, 46(3):200–209.

Becker, A., Schild, U., and Friedrich, C. K. (2018). Tracking independence and merging of prosodic and phonemic processing across infancy. *Developmental Science*, 21(2):1–11.

Benjamin, R. G. and Schwanenflugel, P. J. (2010). Text complexity and oral reading in young readers. *Reading Research Quarterly*, 45(4):388–404.

Bishop, D. V. M. (2003). *Test for Reception of Grammar Version 2*. Department of Experimental Psychology, University of Oxford.

Blythe, H. and Joseph, H. (2011). Children's eye movements during reading. In Liversedge, S., Gilchrist, I., and Everling, S., editors, *The Oxford Handbook of Eye Movements*, chapter 36, pages 643–662. Oxford University Press, United Kingdom.

Blythe, H. I. (2014). Developmental changes in eye movements and visual information encoding associated with learning to read. *Current Directions in Psychological Science*, 23(3):201–207.

Bosse, M.-L. (2015). Learning to read and spell: How children acquire word orthographic knowledge. *Child Development Perspectives*, 9(4):222–226.

Bradley, L. and Bryant, P. (1983). Categorizing sounds and learning to read — a causal connection. *Nature*, 301:419–421.

Cain, K. (2009). Making sense of text: Skills that support text comprehension and its development. *Perspectives on Language and Literacy*, Spring.

Calet, N., Gutiérrez-Palma, N., Simpson, I. C., González, C. M., and Defior, S. (2015). Suprasegmental phonology development and reading acquisition: A longitudinal study. *Scientific Studies of Reading*, 19(1):51–71.

Calet, N., Pérez-Morenilla, M. C., and Santos-Roig, M. D. (2019). Overcoming reading comprehension difficulties through a prosodic reading intervention: A single-case study. *Child Language Teaching and Therapy*, 35(1):75–88.

Carlson, K. (2009). How prosody influences sentence comprehension. *Language and Linguistics Compass*, 3(5):1188–1200.

Carroll, J. M., Bowyer-Crane, C., Duff, F. J., Hulme, C., and Snowling, M. J. (2011). *Developing Language and Literacy: Effective Intervention in the Early Years*. Wiley-Blackwell, Oxford.

Castles, A. and Coltheart, M. (2004). Is there a causal link from phonological awareness to success in learning to read? *Cognition*, 91(1):77–111.

Castles, A. and Nation, K. (2008). Learning to be a good orthographic reader. *Journal of Research in Reading*, 31(1):1–7.

Castles, A., Rastle, K., and Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, 19(1):5–51.

Catts, H. W., Herrera, S., Nielsen, D. C., and Bridges, M. S. (2015). Early prediction of reading comprehension within the simple view framework. *Reading and Writing*, 28(9):1407–1425.

Chan, J. S., Wade-Woolley, L., Heggie, L., and Kirby, J. R. (2020). Understanding prosody and morphology in school-age children's reading. *Reading and Writing*, 33:1295–1324.

Christopher, A., Mehler, J., and Sebastian-Galles, N. (2001). Perception of prosodic boundary correlates by newborn infants. *Infancy*, 2(3):385–394.

Clin, E., Wade-Woolley, L., and Heggie, L. (2009). Prosodic sensitivity and morphological awareness in children's reading. *Journal of Experimental Child Psychology*, 104(2):197–213.

Critten, S., Holliman, A., Hughes, D., Wood, C., Cunnane, H., Pillinger, C., and Deacon, S. (2021). A longitudinal investigation of prosodic sensitivity and emergent literacy. *Reading and Writing*, 34:371–389.

Cromley, J. and Azevedo, R. (2007). Testing and refining the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, 99(2):311–325.

Cromley, J., Snyder-Hogan, L., and Luciw-Dubas, U. (2010). Reading comprehension of scientific text: A domain-specific test of the direct and inferential mediation model of reading comprehension. *Journal of Educational Psychology*, 102(3):687–700.

Cunningham, A. J. and Carroll, J. M. (2015). Early predictors of phonological and morphological awareness and the link with reading: Evidence from children with different patterns of early deficit. *Applied Psycholinguistics*, 36(3):509–531.

Curtin, S. (2010). Young infants encode lexical stress in newly acountered words. *Journal of Experimental Child Psychology*, 105(4):376–385.

Cutler, A. and Carter, D. (1987). The predominance of strong initial syllables in the english vocabulary. *Computer Speech and Language*, 2:133–483.

Cutler, A. and Mehler, J. (1993). The periodicity bias. *Journal of Phonetics*, 21(1):103–108. Phonetic Development.

Deacon, S. H., Holliman, A. J., Dobson, G. J., and Harrison, E. C. J. (2018). Assessing direct contributions of morphological awareness and prosodic sensitivity to children's word reading and reading comprehension. *Scientific Studies of Reading*, 22(6):527–534.

Defior, S., Gutiérrez-Palma, N., and Cano-Marín, M. J. (2012). Prosodic awareness skills and literacy acquisition in spanish. *Journal of Psycholinguistic Research*, 41(4):285–294.

Department for Education (2013). *English programmes of study: Key Stages 1 and 2 framework document*. Department for Education, London, UK. Available at: https://bit.ly/1Vg4Nsj (Accessed: 25th April 2019).

Department for Education (2018). Phonics screening check and key stage 1 assessments: England 2017.

Ehri, L. C. (1995). Phases of development in learning to read by sight. *Journal of Research in Reading*, 18:116–125.

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

Ehri, L. C. (2014). Orthographic mapping in the acquisition of sight word reading, spelling memory, and vocabulary learning. *Scientific Studies of Reading*, 18(1):5–21.

Field, A. P., Field, Z., and Miles, J. (2012). *Discovering statistics using R*. SAGE, London.

Florit, E., Cain, K., and Mason, L. (2019). Going beyond children's single-text comprehension: The role of fundamental and higher–level skills in 4th graders' multiple-document comprehension. *British Journal of Educational Psychology*, 90(2):449–472.

Fodor, J. D. (1998). Learning to parse? *Journal of Psycholinguistic Research*, 27:285–319.

Fodor, J. D. (2002). Prosodic disambiguation in silent reading. In Hirotani, M., editor, *Proceedings of NELS 32*, volume 1, pages 113–132. Amherst: GLSA Publications.

Foorman, B., Herrera, S., Petscher, Y., Mitchell, A., and Truckenmiller, A. (2015). The structure of oral language and reading and their relation to comprehension in kindergarten through grade 2. *Reading and Writing*, 28:655–681.

García, J. R. and Cain, K. (2014). Decoding and reading comprehension: A meta-analysis to identify which reader and assessment characteristics influence the strength of the relationship in English. *Review of Educational Research*, 84(1):74–111.

Gibbon, F. E. and Smyth, H. (2013). Preschool children's performance on profiling elements of prosody in speech-communication (PEPS-C). *Clinical Linguistics and Phonetics*, 27(6-7):428–434.

Gillam, R. B. and Pearson, N. A. (2004). *Test of Narrative Language*. Pro-Ed, Austin, TX.

Godde, E., Bosse, M., and Bailly, G. (2019). A review of reading prosody acquisition and development. *Reading and Writing*, 33:399–426.

González-Trujillo, C. M., Defior, S., and Gutiérrez-Palma, N. (2014). The role of nonspeech rhythm in spanish word reading. *Journal of Research in Reading*, 37(3):316–330.

Goodman, I., Libenson, A., and Wade-Woolley, L. (2010). Sensitivity to linguistic stress, phonological awareness and early reading ability in preschoolers. *Journal of Research in Reading*, 33(2):113–127.

Gough, P. B. and Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7:6–10.

Groen, M. A., Veenendaal, N. J., and Verhoeven, L. (2019). The role of prosody in reading comprehension: Evidence from poor comprehenders. *Journal of Research in Reading*, 42(1):37–57.

Grosjean, P. and Ibanez, F. (2018). *pastecs: Package for Analysis of Space-Time Ecological Series*. R package version 1.3.21.

Gross, J., Millett, A. L., Bartek, B., Bredell, K. H., and Winegard, B. (2014). Evidence for prosody in silent reading. *Reading Research Quarterly*, 49(2):189–208.

Guthrie, J. T., Hoa, L. W., Wigfield, A., Tonks, S. M., Humenick, N. M., and Littles, E. (2006). Reading motivation and reading comprehension growth in the later elementary years. *Contemporary Educational Psychology*, 32:283–313.

Gutiérrez-Palma, N., Defior, S., Jiménez-Fernández, G., Serrano, F., and González-Trujillo, C. M. (2016). Lexical stress awareness and orthographic stress in spanish. *Learning and Individual Differences*, 45:144–150.

Gutiérrez-Palma, N., Raya-García, M., and Palma-Reyes, A. (2009). Detecting stress patterns is related to children's performance on reading tasks. *Applied Psycholinguistics*, 30(1):1–21.

Gutiérrez-Palma, N. and Reyes, A. P. (2007). Stress sensitivity and reading performance in spanish: A study with children. *Journal of Research in Reading*, 30(2):157–168.

Hayes, A. F. (2018). *Introduction to mediation, moderation, and conditional process analysis : a regression-based approach*. New York, N.Y.: Guilford Press.

Hoffman, J. V. (2009). In search of the "Simple View" of reading comprehension. In Israel, S. E. and Duffy, G. G., editors, *Handbook of research on reading comprehension*, pages 54–66. Routledge, New York.

Holliman, A. J. (2016). *Speech rhythm sensitivity in pre-readers: What role does it have in reading acquisition? (Main Report)*. Nuffield Foundation.

Holliman, A. J., Critten, S., Lawrence, T., Harrison, E., Wood, C., and Hughes, D. (2014a). Modeling the relationship between prosodic sensitivity and early literacy. *Reading Research Quarterly*, 49(4):469–482.

Holliman, A. J., Gutiérrez-Palma, N., Critten, S., Wood, C., Cunnane, H., and Pillinger, C. (2017a). Examining the independent contribution of prosodic sensitivity to word reading and spelling in early readers. *Reading and Writing*, 30(3):509–521.

Holliman, A. J., Mundy, I. R., Wade-Woolley, L., Wood, C., and Bird, C. (2017b). Prosodic awareness and children's multisyllabic word reading. *Educational Psychology*, 37(10):1222–1241.

Holliman, A. J., Williams, G. J., Mundy, I. R., Wood, C., Hart, L., and Waldron, S. (2014b). Beginning to disentangle the prosody-literacy relationship: a multi-component measure of prosodic sensitivity. *Reading and Writing*, 27(2):255–266.

Holliman, A. J., Wood, C., and Sheehy, K. (2008). Sensitivity to speech rhythm explains individual differences in reading ability independently of phonological awareness. *British Journal of Developmental Psychology*, 26(3):357–367.

Holliman, A. J., Wood, C., and Sheehy, K. (2010a). The contribution of sensitivity to speech rhythm and non-speech rhythm to early reading development. *Educational Psychology*, 30(3):247–267.

Holliman, A. J., Wood, C., and Sheehy, K. (2010b). Does speech rhythm sensitivity predict children's reading ability 1 year later. *Journal of Educational Psychology*, 102(2):356–366.

Holliman, A. J., Wood, C., and Sheehy, K. (2012). A cross-sectional study of prosodic sensitivity and reading difficulties. *Journal of Research in Reading*, 35(1):32–48.

Holmqvist, K., Nyström, M., Andersoon, R., Dewhurst, R., Halszka, J., and van de Weijer, J. (2011). *Eye Tracking: A Comprehensive Guide to Methods and Measures*. Oxford University Press, Oxford.

Hoover, W. A. and Gough, P. B. (1990). The simple view of reading. *Reading and Writing: An Interdisciplinary Journal*, 2:127–160.

Hoover, W. A. and Tunmer, W. E. (2018). The simple view of reading: Three assessments of its adequacy. *Remedial and Special Education*, 39(5):304–312.

Huey, E. (1908). *The Psychology and Pedagogy of Reading*. Macmillan.

Hulme, C. and Snowling, M. J. (2013). Learning to read: What we know and what we need to understand better. *Child Development Perspectives*, 7(1):1–5.

Jusczyk, P. W., Houston, D. M., and Newsome, M. (1999). The beginnings of word segmentation in english-learning infants. *Cognitive Psychology*, 39(3–4):159–207.

Kaakinen, J. K., Lehtola, A., and Paattilammi, S. (2015). The influence of a reading task on children's eye movements during reading. *Journal of Cognitive Psychology*, 27(5):640–656.

Kent, W. M. V. (2013). *Intonation and Reading Skills in Fourth-Grade Students*. PhD thesis, Wayne State University.

Kilpatrick, D. A. (2012). Not all phonological tests are created equal: considering the practical validity of phonological manipulation versus segmentation. *Communiqué*, 40(6):31–33.

Kim, Y.-S. (2015). Language and cognitive predictors of text comprehension: Evidence from multivariate analysis. *Child Development*, 86(1):128–144.

Kim, Y.-S., Wagner, R. K., and Foster, E. (2011). Relations among oral reading fluency, silent reading fluency, and reading comprehension: A latent variable study of first-grade readers. *Scientific Studies of Reading*, 15(4):338–362. PMID: 21747658.

Kim, Y.-S. G. (2017). Why the Simple View of Reading is not simplistic: Unpacking component skills of reading using a direct and indirect effect model of reading (DIER). *Scientific Studies of Reading*, 21(4):310–333.

Kim, Y.-S. G. and Petscher, Y. (2016). Prosodic sensitivity and reading: An investigation of pathways of relations using a latent variable approach. *Journal of Educational Psychology*, 108(5):630–645.

Kim, Y.-S. G., Vorstius, C., and Radach, R. (2018). Does online comprehension monitoring make a unique contribution to reading comprehension in beginning readers? Evidence from eye movements. *Scientific Studies of Reading*, 22(5):367–383.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2):163–182.

Konza, D. (2014). Teaching reading: Why the "Fab Five" should be the "Big Six". *Austrailian Journal of Teacher Education*, 39(12):153–169.

Kuhn, M. R., Schwanenflugel, P. J., Meisinger, E. B., Levy, B. A., and Rasinski, T. V. (2010). Aligning theory and assessment of reading fluency: Automaticity, prosody, and definitions of fluency. *Reading Research Quarterly*, 45(2):230–251.

Lin, C. Y., Wang, M., Newman, R. S., and Li, C. (2018). The development of stress sensitivity and its contribution to word reading in school-aged children. *Journal of Research in Reading*, 41(2):259–277.

Lochrin, M., Arciuli, J., and Sharma, M. (2015). Assessing the relationship between prosody and reading outcomes in children using the PEPS-C. *Scientific Studies of Reading*, 19(1):72–85.

Lyytinen, H., Shu, H., and Richardson, U. (2015). Predictors of reading skills across languages. In Blavin, E. L. and Naigles, L. R., editors, *The Cambridge Handbook of Reading*, chapter 32, pages 703–723. Cambridge University Press, Cambridge.

Mahmoudzadeh, M., Dehaene-Lambertz, G., Fournier, M., Kongolo, G., Goudjil, S., Dubois, J., Grebe, R., and Wallois, F. (2013). Syllabic discrimination in premature human infants prior to complete formation of cortical layers. *Proceedings of the National Academy of Sciences of the United States of America*, 110(12):4846–4851.

Männel, C., Schaadt, G., Illner, F. K., van der Meer, E., and Friederici, A. D. (2017). Phonological abilities in literacy-impaired children: Brain potentials reveal deficient phoneme discrimination, but intact prosodic processing. *Developmental Cognitive Neuroscience*, 23:14–25.

Marinus, E. and Castles, A. (2015). Precursors to reading: phonological awareness and letter knowledge. In Blavin, E. L. and Naigles, L. R., editors, *The Cambridge Handbook of Reading*, chapter 30, pages 661–680. Cambridge University Press, Cambridge.

McKinlay, A. (2011). *Peabody Picture Vocabulary Test –Third Edition (PPVT-III)*, pages 1072–1072. Springer US, Boston, MA.

Melby-Lervåg, M., Lyster, S., and Hulme, C. (2012). Phonological skills and their role in learning to read: A meta-analytic review. *Psychological Bulletin*, 138(2):322–352.

Miller, B. and O'Donnell, C. (2013). Opening a window into reading development: Eye movements' role within a broader literacy research framework. *School Psychology Review*, 42(2):123–139.

Miller, J. and Schwanenflugel, P. J. (2006). Prosody of syntactically complex sentences in the oral reading of young children. *Journal of Educational Psychology*, 98(4):839–853.

Miller, J. and Schwanenflugel, P. J. (2008). A longitudinal study of the development of reading prosody as a dimension of oral reading fluency in early elementary school children. *Reading Research Quarterly*, 43(4):336–354.

Moats, L. (2010). *Speech to print: Language essentials for teachers. Second Edition.* Brookes Publishing Co., Baltimore.

Moats, L. C. (1999). *Teaching Reading is Rocket Science: What Expert Teachers of Reading Should Know and Be Able to Do*. American Federation of Teachers, Washington, DC.

Moats, L. C. (2020). *Teaching Reading is Rocket Science, 2020: What Expert Teachers of Reading Should Know and Be Able to Do*. American Federation of Teachers, Washington, DC.

Nation, K. (2009). Form-meaning links in the development of visual word recognition. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1536):3665–3674.

Paige, D. D., Rasinski, T. V., and Magpuri-Lavell, T. (2012). Is fluent, expressive reading important for high school readers? *Journal of Adolescent & Adult Literacy*, 56(1):67–76.

Paige, D. D., Rupley, W. H., Smith, G. S., Rasinski, T. V., Nichols, W., and Magpuri-Lavell, T. (2017). Is prosodic reading a strategy for comprehension? *Journal for Educational Research Online*, 9(2):245–275.

Peppé, S. (2015). *Profiling Elements of Prosody in Speech-Communication 2015*.

Peppé, S. and McCann, J. (2003). Assessing intonation and prosody in children with atypical language development: the PEPS-C test and the revised version. *Clinical Linguistics & Phonetics*, 22:363–370.

Perfetti, C. and Adlof, S. (2012). Reading comprehension: a conceptual framework from word meaning to text meaning. In Sabatini, J., Albro, E., and O'Reilly, T., editors, *Measuring up: advances in how to assess reading ability*, pages 3–20. Rowman and Littlefield Education, Lanham.

Perfetti, C. and Hart, L. (2002). The lexical quality hypothesis. In Ludo Verhoeven, C. E. and Reitsma, P., editors, *Precursors of functional literacy*, volume 11, pages 189–213. John Benjamins Publishing, Amsterdam.

Perfetti, C. and Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18(1):22–37.

Peterson, R. A. and Cavanaugh, J. E. (2019). Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. *Journal of Applied Statistics*, pages 1–16.

Petscher, Y., Cabell, S. Q., Catts, H. W., Compton, D. L., Foorman, B. R., Hart, S. A., Lonigan, C. J., Phillips, B. M., Schatschneider, C., Steacy, L. M., Terry, N. P., and Wagner, R. K. (2020). How the science of reading informs 21st-century education. *Reading Research Quarterly*, 55(S1):S267–S282.

Pickering, S. and Gathercole, S. (2001). *Working Memory Test Battery for Children (WMTB-C): Manual*. Pearson.

Preacher, K. J. and Hayes, A. F. (2004). Spss and sas procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers*, 36:717–731.

Quam, C. and Swingley, D. (2012). Development in children's interpretation of pitch cues to emotions. *Child Development*, 83(1):236–250.

Quam, C. and Swingley, D. (2014). Processing of lexical stress cues by young children. *Journal of Experimental Child Psychology*, 123(1):73–89.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rayner, K., Ardoin, S., and Binder, K. (2013). Children's eye movements in reading: A commentary. *School Psychology Review*, 42(2):223–233.

Revelle, W. (2019). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University, Evanston, Illinois. R package version 1.9.12.

Rose, J. (2006). *Independent review of the teaching of early reading: Final report*. Department for Education and Skills Publications, London, UK.

Ryu, C. (2020). *dlookr: Tools for Data Diagnosis, Exploration, Transformation*. R package version 0.3.13.

Schild, U., Becker, A. B. C., and Friedrich, C. K. (2014). Processing of syllable stress is functionally different from phoneme processing and does not profit from literacy acquisition. *Frontiers in Psychology*, 5:1–12.

Schrauben, J. E. (2010). Prosody's contribution to fluency: An examination of the theory of automatic information processing. *Reading Psychology*, 31(1):82–92.

Schroeder, S., Hyönä, J., and Liversedge, S. (2015). Developmental eye-tracking research in reading: Introduction to the special issue. *Journal of Cognitive Psychology*, 27(5):500–510.

Schwanenflugel, P. J. and Benjamin, R. G. (2017). Lexical prosody as an aspect of oral reading fluency. *Reading and Writing*, 30(1):143–162.

Schwanenflugel, P. J. and Klapp, N. F. (2015). *The Psychology of Reading: Theory and Applications*. Guilford Publications.

Seidl, A. (2007). Infants' use and weighting of prosodic cues in clause segmentation. *Journal of Memory and Language*, 57:24–48.

Seymour, P. H. K., Aro, M., Erskine, J. M., and collaboration with COST Action A8 network (2003). Foundation literacy acquisition in european orthographies. *British Journal of Psychology*, 94(2):143–174.

Shapiro, L. R., Carroll, J. M., and Solity, J. E. (2013). Seperating the influences of prereading skills on early word and nonword reading. *Journal of Experimental Child Psychology*, 116(2):278–295.

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

Share, D. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, 72:95–129.

Singh, L., Hui, T. J., Chan, C., and Golinkoff, R. M. (2014). Influences of vowel and tone variation on emergent word knowledge: a cross-linguistic investigation. *Developmental Science*, 17:94–109.

Snow, P. C. (2020). Solar: The science of language and reading. *Child Language Teaching and Therapy*, pages 1–12. Preprint on webpage at https://journals.sagepub.com/doi/10.1177/0265659020947817.

Snowling, M. J. and Hulme, C. (2005). *The Science of Reading: A Handbook*. Blackwell Publishing Ltd.

Snowling, M. J., Stothard, S. E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., Nation, K., and Hulme, C. (2011). *York Assessment of Reading for Comprehension Passage Reading - Second Edition*. GL Assessment, Brentford, UK.

Soderstrom, M., Nelson, D., and Jusczyk, P. (2003). The prosodic bootstrapping of phrases: Evidence from prelinguistic students. *Journal of Memory and Language*, 49(2):249–267.

Stainthorp, R. (2020). A national intervention in teaching phonics: A case study from england. *The Educational and Developmental Psychologist*, 37(2):1–9.

Stuart, M., Stainthorp, R., and Snowling, M. (2008). Literacy as a complex activity: Deconstructing the Simple View of Reading. *Literacy*, 42(2):59–66.

Thomson, J. M. and Jarmulowicz, L. (2016). *Linguistic Rhythm and Literacy*. John Benjamins Publishing Company, Philadelphia.

Tingley, D., Yamamoto, T., Hirose, K., Keele, L., and Imai, K. (2014). mediation: R package for causal mediation analysis. *Journal of Statistical Software*, 59(5):1–38.

Torgesen, J. K., Wagner, R. K., and Rashotte, C. A. (2012). *Test of Word Reading Efficiency - Second Edition*. Pro-Ed, Austin, TX.

Treiman, R. and Kessler, B. (2005). Writings systems and spelling development. In Snowling, M. and Hulme, C., editors, *The Science of Reading: A Handbook*, pages 120–134. Blackwell Publishing, Oxford.

Trueswell, J. C., Sekerina, I., Hill, N. M., and Logrip, M. L. (1999). The kindergarten-path effect: studying on-line sentence processing in young children. *Cognition*, 73(2):89–134.

Tunmer, W. A. and Hoover, W. A. (2019). The cognitive foundations of learning to read: a framework for preventing and remediating reading difficulties. *Australian Journal of Learning Difficulties*, 24(1):75–93.

Uppstad, P. H. and Solheim, O. J. (2011). Code and comprehension in written language - considering limitations to the simple view of reading. *L-1-Educational Studies in Language and Literature*, 11:159–174.

Valle, A., Binder, K., Walsh, C., and K.E., B. (2013). Eye movements, prosody, and word frequency among average- and high-skilled second-grade readers. *School Psychology Review*, 42(2):171–190.

Van Orden, G. C. and Kloos, H. (2005). The question of phonology and reading. In Snowling, M. J. and Hulme, C., editors, *The Science of Reading: A Handbook*, chapter 4, pages 61–81. Blackwell Publishing, Oxford.

Van Wingerden, E., Segers, E., Van Balkom, H., and Verhoeven, L. (2018). Cognitive constraints on the simple view of reading: A longitudinal study in children with intellectual disabilities. *Scientific Studies of Reading*, 22:321–334.

Veenendaal, N. J., Groen, M. A., and Verhoeven, L. (2014). The role of speech prosody and text reading prosody in children's reading comprehension. *British Journal of Educational Psychology*, 84:521–536.

Veenendaal, N. J., Groen, M. A., and Verhoeven, L. (2015). What oral text reading fluency can reveal about reading comprehension. *Journal of Research in Reading*, 38(3):213–225.

Veenendaal, N. J., Groen, M. A., and Verhoeven, L. (2016). The contribution of segmental and suprasegmental phonology to reading comprehension. *Reading Research Quarterly*, 51(1):55–66.

Verhoeven, L. and van Leeuwe, J. (2008). Prediction of the development of reading comprehension: a longitudinal study. *Applied Cognitive Psychology*, 22(3):407–423.

Wade-Woolley, L. (2016). Prosodic and phonemic awareness in childrens reading of long and short words. *Reading and Writing*, 29:371–382.

Wade-Woolley, L. and Heggie, L. (2015). Implicit knowledge of word stress and derivational morphology guides readers' decoding of multisyllabic words. *Scientific Studies of Reading*, 19:21–30.

Wagner, R., Torgeson, J., Rashotte, C., and Pearson, N. A. (2013). *Comprehensive Test of Phonological Processing - Second Edition*. Pro-Ed, Austin, TX.

Wagner, R. K. and Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 101(2):192–212.

Wei, T. and Simko, V. (2017). *R package "corrplot": Visualization of a Correlation Matrix*. (Version 0.84).

Wennerstrom, A. (2001). *The Music of Everyday Speech: Prosody and Discourse Analysis*. Oxford University Press, New York.

Whalley, K. and Hansen, J. (2006). The role of prosodic sensitivity in children's reading development. *Journal of Research in Reading*, 29(3):288–303.

Whalley, K. M. (2017). *The Role of Prosodic Skills in Reading Comprehension*. PhD thesis, Queensland University of Technology.

Wiederholt, J. and Bryant, B. R. (2001). *Gray oral reading test (4th ed.)*. Pro-Ed, Austin, TX.

Wigfield, A., Gladstone, J. R., and Turci, L. (2016). Beyond cognition: Reading motivation and reading comprehension. *Child Development Perspectives*, 10(3):190–195.

Wolf, M., Muijselaar, M., Boonstra, A., and de Bree, E. (2019). The relationship between reading and listening comprehension: shared and modality-specific components. *Reading and Writing*, 32:1747–1767.

Wood, C. (2006). Metrical stress sensitivity in young children and its relationship to phonological awareness and reading. *Journal of Research in Reading*, 29(3):270–287.

Wood, C. and Terrell, C. (1998). Poor readers' ability to detect speech rhythm and perceive rapid speech. *British Journal of Developmental Psychology*, 16(3):397–413.

Wood, C., Wade-Woolley, L., and Holliman, A. (2009). Phonological awareness: Beyond phonemes. In Wood, C. and Connelly, V., editors, *Contemporary perspectives on reading and spelling*, pages 7–23. Routledge, London.

Ziegler, J. C. and Goswami, U. (2005). Reading acquisition, developmental dyslexia, and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin*, 131(1):3–29.

Ziegler, J. C. and Goswami, U. (2006). Becoming literate in different languages: Similar problems, different solutions. *Developmental Science*, 9(5):429–453.