The Vocabulary Gap in Pre-School Children: a Corpus-Based Analysis and Exploration of Audio Intervention via Action Research

Hannah Phelan

Nottingham Trent University

Abstract

Set against a context of robust existing research showing both differences between children's early vocabularies (the 'vocabulary gap'), and associations between smaller vocabularies and later quality of life indicators, Phase One of this thesis sought to investigate the vocabulary gap in terms of its content and risk factors. Phase Two then looked to explore an intervention format which could support children with smaller vocabularies. To these ends, and using predictions based on age of acquisition research, a corpus of children's talk was analysed to explore the differences in children's speech across a range of vocabulary sizes and against a number of linguistic (e.g., word frequency; lexical category) and demographic (e.g., income; number of siblings) variables. The aim of this analysis was to ascertain which linguistic variables were predictive of vocabulary (therefore shedding some light on the content of the vocabulary gap), as well as which demographic variables were predictive of vocabulary (therefore providing useful tools to identify children in need of extra support). The most robust result emerging from this analysis was the importance of verbs, in that children who use more verbs in their speech tend to have larger vocabularies overall, with further analysis suggesting that this might be because verbs facilitate the acquisition of additional vocabulary. In Phase Two, parent-led interventions for the pre-school age range were reviewed, and limitations noted. The audiobook was argued to be a format capable of circumventing some of these limitations, and thus Action Research methodology was employed, involving reflective input from both parents and the researcher, to explore the potential of and develop a simple audiobook intervention. Phase Two built on Phase One by incorporating different levels of verb diversity within the audio scripts. Findings suggested that audiobooks have the potential to engage parents and children, in terms of their convenience, accessibility, and ability to stimulate parent-child interaction, although the extent of this potential will depend on many factors, including a family's existing routines. Limitations and wider implications are discussed, including an outline proposal for future research in the final chapter.

Contents

Chapter 1: Introduction and Theoretical Perspectives

A notable proportion of children are starting school without the language skills expected for a child of their age (Finnegan & Warren, 2015). Teachers report up to 40% of children are lacking the necessary vocabulary to access learning (Harley et al., 2018) and, in 2015, almost 20% of 5-year-olds were unable to meet the language-related Early Learning Goals at the end of Reception year, with this result being twice as likely for those on free school meals (Law et al., 2017a). Early language ability, including vocabulary knowledge, has been robustly linked to numerous quality of life indicators (e.g., Nightingale, 2019), and thus difficulties are important to address. Not all children experience language difficulties however, and the difference between those with largest and smallest vocabularies is often termed the 'language gap' or 'vocabulary gap.' Researchers studying the vocabulary gap are usually concerned with differences in vocabulary which are due to environmental input, rather than specific developmental or medical conditions, although these groups are not always excluded (Finnegan & Warren, 2015).

Language is concerned with the words we use – vocabulary – and how those words are put together – grammar (Harley, 2013). Thus, the development of language skills is concerned with acquiring aptitudes along these two dimensions. Language is about communicating with others, but it is also about thinking and expressing emotion (Crystal, 2010), and is generally considered one of the hallmarks of our species (Harley, 2013). Despite the terms *language* and *literacy* often being used in conjunction with one another, they are concerned with different (albeit related) aptitudes, with literacy skills associated with the later developmental stages of reading and writing, including phonological awareness and print awareness. Language skills are an important predecessor to literacy skills, in that oral language skills feed into both phonemic awareness and reading comprehension, which arise from having heard, practised and understood the meaning of words. Alongside literacy development, vocabulary knowledge has been linked to a number of other important outcomes both in the short and long term. In practice, children who arrive at school without the expected level of vocabulary for their age are likely to experience difficulty following instructions, articulating their needs and feelings, and be lacking in ideas for socialising and play (Harley et al., 2018). Beyond school, the words children know will help them to comprehend humour (Purser et al., 2020), develop their self-esteem and confidence (Harley et al., 2018), and nurture important skills for future employment such as the ability to resolve conflicts and persuade others (Gascoigne & Gross, 2017). Early language difficulties have also been linked to reduced social and emotional well-being (e.g., an increased likelihood of psychiatric disorders; Beitchman et al., 2001) youth offending (Royal College of Speech & Language Therapists, 2012), and even life expectancy (Gilbert et al., 2018), as well as later academic results (Snowling et al., 2011) and future employment (Nightingale, 2019).

As noted above, language is imperative for helping children achieve academic and exam success, currently a cornerstone of the UK education system. In a survey of 473 primary school and 840 secondary school teachers (Harley et al., 2018), almost all primary school teachers felt that children with a limited vocabulary made slower progress in reading and writing, and 80% of secondary teachers felt it would be extremely challenging for pupils with limited vocabulary to read national test papers. Further suggestions for how vocabulary difficulties may impact academic achievement included difficulties in working independently (75% of teachers agreed), difficulties following what was going on in class (77% of teachers agreed), as well as wider negative impacts on self-esteem (80%

agreed), behaviour (65% agreed) and attendance (34% agreed) – all of which provide further explanation for the link between vocabulary and later academic achievement. Evidence of attainment at different stages is in agreement with teachers and further supports the link between language, and specifically vocabulary and academic achievement: 88% of 4 and 5-year-old children with a language disorder did not make expected academic progress (Norbury et al., 2016), and only 12% of children with a Speech Language and Communication Need (as their main need) achieved the expected standard in reading, writing and maths by the end of primary school, compared to 53% of all pupils (Gascoigne & Gross, 2017). Furthermore, vocabulary tested at age 13-14 years predicted GCSE results in both English and Maths (Spencer et al., 2017).

Challenging lower-than-expected language (and literacy) levels throughout childhood has thus become a prominent item on the political agenda, with recent pledges to ´close the gap´ from both the Education Department and the business sector, including a £5 million fund entrusted to the Education Endowment Foundation in 2018 to develop and trial interventions aiming to improve children's language and literacy in the pre-school years (Department for Education, 2018), as well as a recent post-election commitment from the incoming Education Secretary to fund the Nuffield Early Language Intervention for its fifth year. As yet however, there has been no panacea. The problem is exacerbated by difficulties in identifying those who are most in need of intervention: the high degree of variability in children's language trajectories in the early years means that pinpointing those who are behind versus those who are on track is not necessarily a straight-forward task. One common approach to targeting support is to use social background (e.g., Pupil Premium funding), and yet to use this factor alone is also problematic. Whilst a link appears to exist between socio-economic (SES) background and children's language outcomes (Hart & Risley, 1995; Hoff, 2003; Rowe 2017), difficulties can be attributed to a range of factors and affect children across the SES spectrum (Law et al., 2017a) - thus illustrating the complexity of the task at hand.

The remainder of this chapter will begin by exploring the prevalence of language difficulties amongst children in the UK, and the complexities inherent in identifying those children who may be at greater risk. It will then consider some of the major theoretical approaches to language acquisition (in particular, word acquisition) set against the context of the stages of children's language acquisition.

Exploring the Prevalence of Language Difficulties

Speech, Language and Communication Needs (SLCN) is an umbrella term made familiar by the Bercow Review (2008) and used frequently by schools and other professionals: it can be interpreted broadly or narrowly. The Bercow Review (2008) utilises SLCN as a broad category, referring to any child with any language and communication need, including the 7% of all children thought to have a developmental language disorder (DLD, previously known as Specific Language Impairment), the 3% of all children with a Language Disorder (LD) linked to another condition, such as cerebral palsy or hearing loss, plus any other child who may be experiencing difficulties or delays, including speech sound disorders, late talkers, or those experiencing language delay due to environmental factors. It is this very latter group of children experiencing language delay that are of particular interest when studying the vocabulary gap; children who have no specific difficulty associated with their language development (such as DLD) but have nonetheless fallen behind their peers and the expected developmental milestones for their age range (Finnegan & Warren, 2015, p.2).

Ascertaining the prevalence of language delay remains elusive however, and national figures that refer to SLCN should be treated with caution: if being used in its broadest sense, the term SLCN will include this group (The Bercow Review, 2008), however if used in a narrower sense (e.g., schools' data may record only those children with a specific DLD or LD as having a SLCN), then it will not. Research investigating the prevalence of language needs tends not to distinguish between those with DLDs, LDs, language delay or any other type of language difficulty. Indeed, even studies that declare a specific interest in children with a language delay (as opposed to any other kind of language difficulty), as was the case in Warren and Finnegan (2015), neither of the datasets used excluded children with DLDs and LDs, leading to the conclusion that any child who did not meet the required standard could have done so for a multitude of reasons, not necessarily a language delay specific to environmental circumstances.

Ambiguity is still apparent even when the broadest interpretation of SLCN is considered, and no single agreed ballpark statistic exists regarding how widespread the level of need is. For example, one study (using teacher report across a range of age groups) declared that 40% of children lack the necessary vocabulary to access learning (Harley et al., 2018); another found the prevalence of SLCN in children starting school in areas of social disadvantage to be as high as 50% when children's language was tested in the first term of Reception class (Locke et al., 2002); and yet another found that 1 in 4 children (rising to 1 in 3 for those on FSM) were unable to meet the language-related Early Learning Goal (Finnegan & Warren, 2015). Indeed, using the same data but from the following year, Law et al. (2017a) found that 1 in 5 children did not meet the same goal, with this result being twice as likely for those on FSM. These statistics illustrate that the exact prevalence of SLCN in their broadest sense is fluid in nature and difficult to ascertain, which necessarily has knock-on effects when considering more specific categories of SLCN, in this case language delay attributed to environmental factors. Nonetheless, if, according to the statistics above, between 20-50% of all children have some sort of SLCN, and 10% of children have a more specific DLD or LD associated with another condition (Norbury et al., 2016), then a sensible estimate for other SLCN (with 'other' including those who are experiencing delays due to environmental factors) would be between 10 – 40%.

Despite the disparity regarding exact prevalence rates of SLCN in the UK, there appears to be a consensus amongst different parties that children's SLCN are needs which exist and require addressing. Commissioned by the Government, The Bercow Review (2008) identified a number of key themes regarding speech and language provision across the country, including the importance of early identification and intervention, as well as joined-up service provision (which was found to be highly variable and inequitable). Subsequent years saw several positive changes initiated, including a move to bring language and communication into a stronger position within the Early Years and Ofsted Inspection Framework, and a renewed focus on language at the 2 to 2-and-a-half year Health Visitor check. Yet there is still much left to accomplish: many teachers and other professionals working with younger children may not possess the knowledge required to identify early indicators of SLCN (for example difficulties with early use of pointing and gesture), and many language screening tools are not sensitive enough to pick up on these indicators (Bercow 10 years on, 2018). Oral language skills are checked universally only twice (in the UK) before the age of 11, at age 2 and age 5, leading to a call for them to take a more prominent position in the curriculum across all age groups, including more frequent assessments, and a gold standard benchmark for identifying language difficulties (Dockrell, 2019).

Identifying those at Risk

The Bercow Review (2008) indicated that early identification and intervention was a key area of improvement, and yet identifying which children require support with their language development can be a complex task. Where a child has a LD associated with another known condition, then this condition may serve as an early red flag for any additional needs relating to language development – however for other types of language difficulty, there is likely to be no obvious red flag of this kind, and children who should develop language skills at the same rate as their peers sometimes do not. With vocabulary gaps emerging as early as 18 months (Fernald et al., 2013), identification of those experiencing delays should ideally take place before formal schooling begins in order to have maximum impact (Heymann et al., 2020), and yet during this period the onus of recognising language delays tends to fall on the parents, who might not have the necessary knowledge or experience to identify additional language needs in their children. To exacerbate the complexity of identifying those in need, universal screening instruments to assess pre-schoolers' language abilities are limited in their reliability, particularly when complicated by the natural variance in children's language trajectories (Siu, 2015). Indeed, some children with initial delays seem to catch up without obvious intervention; whilst others have difficulties which persist, and a minority may start off well but develop additional needs later on.

Some commentators have suggested a more integrated approach to identifying children with or at risk of language delay or difficulty. Law et al. (2017a) proposed an integration of child, family and parenting factors to establish risk, allowing a gradient response to a continuum of need, rather than a simple cut off point concerning those who require support and those who do not. Approaches which focus on identifying children with multiple vulnerabilities, such as social and emotional difficulties alongside language difficulties (Girard et al., 2016), or speech and language difficulties (Hayiou-Thomas et al., 2016), have also been proposed. Furthermore, approaches that allow for the monitoring of progress over a period of time might be helpful in mitigating the complications presented by natural variation in child language acquisition, in that different assessment points would pinpoint the delays which were persisting and in need of support. The latter approach has been partially implemented by the adoption of the Ages and Stages Questionnaire (ASQ), part of the 2- to 2-and-a-half-year toddler health check carried out by Health Visitors. Indeed, this group of professionals appear to be particularly well placed to facilitate such assessments, in that the mandatory health checks they perform between 0-5 years may be the only professional contact a child receives before they attend formal schooling. A recent collaborative initiative commissioned by Public Health England aims to develop a new early assessment tool, incorporating more of the red flags mentioned above (Law et al., 2017a), and train Health Visitors to deliver such a tool alongside the ASQ. Trials were due to go ahead in 2019/20 in five Local Authorities across the country but were delayed due to the COVID-19 outbreak.

Another group of people (due to their training, experience and contact with children before school-age) who are potentially well-placed to identify children who may have or be at risk of early language difficulties, as well as provide quality verbal input to children, are those who work in early years childcare settings. Indeed, one of the Government's key aims in introducing the universal entitlement of 15 hours of free childcare for 3 and 4-year-olds back (and some 2-year-olds where

families were in receipt of specific benefits) in 1998 was to support child development. Attendance at early childcare settings has been consistently linked with better language outcomes for children (e.g., Davies et al., 2021), particularly for children from families where parents have lower education levels (Becker, 2011), although this is partly dependent on the quality of childcare provided (Melhuish, 2016).

The more recent entitlement of 30 hours of free childcare introduced in 2017 has an additional aim of supporting maternal employment, although is only available where both parents are working and earning over a minimum threshold, hence does not benefit children from the poorest families. Despite this, 30% of local authorities felt that the roll out of 30 hours would have a positive (narrowing) effect on the attainment gap between disadvantaged children and their peers (Coleman & Cottell, 2019), but benefits may be moderated by poor uptake due to the complexity of the system, extra hidden costs (Coleman & Cottell, 2019), and lack of available places, particularly good quality ones (Pollard et al, 2023).

Summary: Prevalence and Identifying those at Risk

Evidence suggests that a notable proportion of children are starting school without the language skills expected for a child of their age (Finnegan & Warren, 2015; Harley et al., 2018; Law et al., 2017a), including smaller or 'delayed' vocabularies. This has the potential to adversely affect children's outcomes in the short-term (with respect to school experience and academic success) and the long-term (with respect to job prospects and general well-being). Identifying children in need or at risk of language delay or difficulties in general is not straight forward before formal school begins, as there are few official language checks for the under 5s. Currently, identifying potential difficulties with language falls mostly to the parent, or, particularly after the introduction of 30 hours free childcare, to early years practitioners.

Theoretical Approaches to the Study of Language Acquisition

It is important to consider perspectives on how children learn language, as this has implications for what sort of intervention might support a child experiencing language difficulties. In simplistic terms, theories of general language acquisition (accounting for vocabulary and grammar) have traditionally split along the nature-nurture divide, with nativist approaches arguing for a more innate mechanism, supported by research that focuses on the biological influences on language, and constructivist and behaviourist theories focusing primarily on the language environment.

Nativist Theories

Chomsky (1957) argued for an innate language capability in the form of the Language Acquisition Device, and a Universal Grammar shared across all languages, evidenced by the fact that all children, regardless of language being learned, tended to hit the same milestones at approximately the same time. He asserted that learning language through imitation alone could not account for a child's ability to produce novel sentences – the aptly named poverty of the stimulus argument – and parents were inconsistent in both the way they modelled language and in their use of negative reinforcement. An important component of Chomsky's theory was the idea of the Language Acquisition Device (LAD), an innate, domain specific mechanism within the brain that set a number of principles and parameters within which language acquisition was constrained. Thus, children were

born with an innate understanding of grammatical rules which allowed them to acquire vocabulary quickly and efficiently, and this explained why they could recognise sentences that were grammatically correct but did not make sense (Chomsky, 1957). The idea that syntactic knowledge was too complex to be gleaned exclusively from input alone (poverty of the stimulus) has evoked much controversy and challenge, the complexities of which are beyond the scope of this thesis (see e.g., Pullum & Scholz, 2008, and Behme & Deacon, 2008). Chomsky's ideas were (and still are) highly influential, but were challenged by opponents who advocated for the greater importance of environment factors in language learning, as discussed below.

Biological Explanations

FMRI and lesion studies have implicated multiple areas of the brain to be involved in language learning, evidence that appears to support the idea of an innate language capability in humans. For most, the left hemisphere appears to be dominant for language (Harley, 2013), although this is not always the case. Specific areas of the brain have been associated with specific language skills, for example Broca's area (speech production) and Wernicke's area (comprehension), illustrated by the fact that when damage occurs to these particular locations, specific and predictable difficulties occur, for example Broca's aphasia (Lau et al., 2021). On the other hand, Farah (1994) challenges the assumption of 'locality,' and the modularity hypothesis in particular (whereby modules are specialised functional components, with interaction of specialised parts producing the behaviour), arguing that even simple tasks involve a network of several different components at once. Modularity is often used to explain the origin of uneven cognitive profiles in developmental disorders, such as Williams Syndrome (Bellugi et al., 2000), yet it has been counter-argued that caution should be employed when tempted to over-state such uneven cognitive profiles, as from a neuroconstructivist position, if modules are a product of development, and development is faulty in some way, it is difficult to see how only one module would be impaired (Thomas et al., 2014). This is particularly true of Williams Syndrome; Brock et al. (2007) found that the strengths and difficulties with respect to language and visuo-spatial skills to be uneven only when the very extremes were measured. In addition, what appeared to be linguistic competence with respect to figurative language did not necessarily reflect a normal underlying process upon further investigation (Thomas et al., 2014). This casts doubt on the notion that language learning is completely modular and suggests that a complex integration of multiple systems is more likely, although either scenario supports the notion that humans possess some level of innateness for language.

Variability in word learning can be explained to some degree by genetic factors (Stromswold, 2001), particularly via cognitive skills such as short-term memory (Gathercole et al., 1997), processing efficiency (Fernald et al., 2006) and attention span (Jones & Rowland, 2017). Twin and adoption studies have provided further support for theories arguing for the innateness of language. In a review of the evidence, Stromswold (2001) suggested that genetic factors played a role in both the rate of language acquisition and the eventual level of linguistic proficiency gained. For vocabulary in particular, the reviewed studies showed between 25%-50% of the variance in the spoken language of normal twins was due to heritable factors, and most likely it was many genes acting in combination with environmental factors (Plomin & Dale, 2000). Puglisi et al., (2017) argued for the influence of a passive gene-environment correlation whereby parents who can read well not only pass this on genetically, but are also more likely to provide a stimulating home learning environment. The fact that storybook

exposure was not a significant predictor of children's outcomes once maternal language skill had been accounted for suggested a genetically mediated explanation for the link between the home learning environment and children's literacy skills, although maternal education could be another plausible driver of effects, and other studies have found the association between storybooks and language outcomes to hold, even once maternal skills have been accounted for (Lonigan & Whitehurst, 1998). These somewhat contradictory findings can perhaps be unified by the notion that whilst the emphasis of a particular theory may veer towards nature or nurture, a comprehensive theory of language acquisition needs to account, to varying degrees, for both.

Behaviourist and Constructivist Theories

Early behaviourist theories (Skinner, 1957) posited that children learned language via imitation and associating objects with words, with positive reinforcement causing them to realise their communicative importance, and thereby increasing their number of correct attempts. In this way, verbal behaviour worked in much the same way as other behaviour. Later on, those arguing for a more constructivist paradigm (including functional, usage-based, social-pragmatic and social-interactionist approaches, e.g., Tomasello, 2005; Bybee & Beckner, 2012; Ambridge & Lieven, 2015) believed that children used a general cognitive mechanism to learn language, rather than a language-specific one, and constructed knowledge gradually, dependent on the language input they received. Whilst similar to earlier behaviourist approaches in terms of the importance placed on environmental factors, a child constructing their own knowledge from the environment differs from the behaviourist viewpoint of the child as a passive learner, whose behaviour follows the guide of positive and negative reinforcement provided by the adult.

Summary: Theoretical Approaches to the Study of Language Acquisition

Evidence appears to be mixed in support of one or other of the two main and current approaches: nativism or constructivism, with polarisation more apparent regarding the question of grammar than word learning. Studies showing the presence of prosodic features of adult speech in their infant's babbling (Mampe et al., 2009), the different nature of deaf babbling (Moeller et al., 2007), and the effects of the quantity and quality of verbal input on a child's language development (Hart & Risley, 1995) have all been used to evidence the importance of an environmental component in language acquisition at this stage. At the other end of the spectrum, Chomsky claimed that when an animal and an infant were both exposed to an identical verbal environment, only the infant would develop language, and this in itself was compelling evidence for the innateness of language (Chomsky, 1978), particularly when combined with the twin/adoption studies (Stomswold, 2001) and the acknowledged role of the brain in language acquisition. As already noted, it is highly likely that both nature and nurture play important and interrelated roles in how children learn language (Shanker & Taylor, 2001, p.51). This can be illustrated to some extent by the case of Genie (Curtiss, 2014), whose abusive upbringing was devoid of verbal input. Genie was able to acquire vocabulary and basic syntax later on, but not more advanced syntactic abilities, suggesting that both environmental input and critical or sensitive periods for language development are both important.

Stages of Language Development

Referring to 'stages' of language acquisition and presenting the following information in a broadly linear fashion infers a neat and orderly progression through developmental milestones, which is somewhat misleading. Whilst it is true that there is a co-dependency between many aspects or stages of language acquisition, with some stages seeming to logically precede others, the process is not a fully linear one - indeed evidence suggests that many aspects of language, for example babbling and perception of speech sounds, are 'mutually enforcing' (Guo et al., 2023, p.74) rather than developing separately, in order.

Newborns and Babbling

Whilst language acquisition refers to the process of learning to understand and use words and grammar, the process of building the prerequisite skills for acquiring language begins well before a child can vocalise their first words. Indeed, there appears to be a mechanism for learning language at the pre-natal stage, with studies using a non-nutritive sucking paradigm indicating that newborns prefer both their own mother's voice (DeCasper & Fifer, 1980) and a familiar passage that had been read to them prenatally (DeCasper & Spence, 1986).

Canonical babbling, whereby infants repeat adult-like, consonant and vowel-containing syllables from around 6 months of age, is another important precursor to acquiring language (McGillion et al., 2017), with the quantity and complexity of babbling linked with later vocabulary levels (Morgan & Wren, 2018) and emerging simultaneously with understanding: children begin to link common words and their referents at this age (Bergelson & Swingley, 2012). Both the intonation of a baby's crying (Mampe et al., 2013) and babbling (Kimbrough Oller, 2010) reflect their native language in terms of prosody and phonological patterns, indicating a susceptibility to environmental input from a very early stage of development.

Phonological Development and Speech Production

Another important aspect of language development for both comprehension and production is phonological development - how children learn to understand, use and contrast between the sounds of their native language. Similar to other topics of language development, nativist theories suggest that these sounds are based on an innate set of distinctive features on which all the world's languages are based, whereas more constructive accounts suggest that children acquire these phonological contrasts throughout childhood from the sounds to which they are exposed. They then learn the relationships between the sounds which they hear and those which they themselves are able to make via their vocal tract (including their lungs, vocal cords, oral and nasal cavities and articulators), something which they begin to do around at 7-9 months old (Ambridge & Lieven, 2011). Various researchers (e.g., Jusczyk, 2000) have noted that phonological development is likely to be interwoven with word segmentation (discussed in the section below), and thus despite infants being able to perceive speech sounds from early on, they do not necessarily need to possess a 'fully segmental phonology in advance of speech production' (Ambridge & Lieven, 2011, p.57). The tendency of infants' early speech to differ from that of adults' (e.g., syllable deletions or reduplications of syllables) suggests that both phonological and lexical knowledge, as well as articulation possibilities, continue to influence the words a child produces throughout childhood.

Segmenting from the Speech Stream

Most words spoken to children are done so in multi-word utterances, not as single words (Brent & Siskind, 2001), and hence in order to begin to understand words, children first need to decide which groups of sounds constitute words (i.e., segmenting from the speech stream), a task made more difficult by the fact that words tend to have different lengths and stress patterns, and are subjected to co-articulation when expressed. The question of how children achieve this is disputed, with different explanations, including those focusing on stress patterns and statistical probabilities, made more complex by the fact that children appear to be able to segment words before they form part of their receptive vocabularies (Aslin et al., 1996).

Infants appear to be able to segment some words more easily than others, possibly by employing knowledge of the familiar stress patterns of their language. Using a habituation paradigm, Jusczyk et al. (1999) found that 7-and-a-half-month-olds were able to segment target words with a strong-weak (trochaic) stress pattern (as dominant in the English language), such as *kingdom,* but were unable to detect words with a weak-strong (iambic) stress pattern, such as *guitar*, suggesting that a child's default position was to interpret a strong syllable as the onset of a word. This was not the default position for all children however, with Polka and Sundara (2003) finding the opposite to be true for French-Canadian 8-month-olds, where iambic stress patterns are the norm. Both studies support the idea that children make use of dominant stress patterns via the influence of their native language to help them make sense of the spoken language that they hear. On the other hand, Floccia et al. (2016) reviewed thirteen similar studies of British-English 8- to 10-and-a-half-month-olds, and found only one to replicate the findings in Jusczyk et al. (1999), which concerned American-English children. These differing results were attributed to the different characteristics of infant-directed speech (IDS), with American IDS containing greater prosodic differences between target words, and thereby boosting the salience of the target word (DePaolis et al., 2010), rather than very slight methodological nuances. Indeed, this explanation is supported by the existence of an unexplained vocabulary gap between American and English toddlers (Hamilton et al., 2000), with American children tending to have larger vocabularies in the early years. Only one British study reviewed by Floccia et al. (2016) found evidence of segmentation (Mason-Apps et al., 2011), using older children and an exaggerated style of IDS.

Another way to explain how children segment words is through statistical learning, which describes the process whereby a person can notice and extract structural information from the environment and then use it to define word boundaries (Saffran et al., 1996). In the context of language, statistical learning can take place at phoneme, syllable or word level, and include relative frequencies and transitional probabilities, the latter of which refers to the probability of one sound sequence occurring in relation to an adjacent one (Romberg & Saffran, 2010). Using the same habituation paradigm as Jusczyk et al. (1999), Saffran et al. (1996) were able to demonstrate that 8-month-olds could utilise these transitional probabilities between neighbouring speech sounds to segment nonsense words to which they had been exposed to a short time earlier via a two-minute monotone delivery devoid of stress patterns and other prosodic cues. Marcus et al. (1999) went further by showing that 7-month-olds were not only able to track transitional probabilities within the original stimulus presented, but learn rules and generalise these rules beyond that original stimulus. A later meta-analysis (Rabagliati et al., 2019) revealed that this ability for infants to learn rules was

particularly strong with respect to speech-like stimuli (when compared to other modalities), although it remains unclear how this learning varies across ages, language backgrounds and methodologies: discrepancies which a currently in progress ManyBabies collaborative project on rule learning aims to resolve (Visser et al., pre-print).

The experimental conditions of the aforementioned studies are not necessarily reflective of real-life language exposure (as conceded by Saffran et al., 1999), and yet the skill of extracting statistical information from limited exposure, and potentially applying that learning in other contexts, would presumably work alongside other cues in more naturalistic settings. The idea that children use multiple cues to aid learning, with different cues being more important at different developmental stages (Jones et al., 2021) can be considered in parallel with research showing that infants benefit differently from quantity or diversity of input dependent on their age (Rowe, 2012), although the exact nature of how such cues are weighted, and whether this is dependent on age or situation (or both) remains elusive. In a later review, Romberg and Saffran (2010) suggested that infants made use of all cues available to them, including some sort of bootstrapping, with transitional probabilities used first (Thiesson & Saffran, 2003), then, as the lexicon develops, a move towards using stress patterns (Johnson & Jusczyk, 2001), although, as discussed, stress patterns appear to be more influential in American-English rather than British-English, where IDS has a more exaggerated prosody (DePaolis et al., 2010). Jusczyk et al. (1999) found that once infants were 10-and-a-half months old, they showed no preference for the dominant stress pattern of their native language, perhaps suggesting a return to statistical patterns as the dominant cue, but this time at word, syllable and phoneme level. As a child's vocabulary expands, word segmentation becomes easier and faster as existing lexical and sub-lexical knowledge can act as both an anchor point to surrounding unfamiliar words (Bortfield et al., 2005), in addition to allowing similar sounding words to be segmented more easily (Alvater-Mackenson & Nivedita, 2013). Indeed, Jones et al. (2021) found that children's word-learning was highly dependent on prior sub-lexical knowledge from known word types, with phonological knowledge acquisition involving a gradual accumulation of larger and larger chunks of phoneme sequences.

Development of Non-verbal Abilities

The development of pre-linguistic non-verbal abilities, such as pointing, turn-taking, eye gaze and other symbolic gestures have been found to be an important precursor in that they have been strongly associated with later rates of language development. This is particularly true of declarative gestures, such as pointing (Iverson & Goldin-Meadow, 2005), but also symbolic gestures, such as flapping the arms to imitate a bird's flying (Goodwyn et al., 2000). From around 10 months of age, infants are able to engage in triadic joint attention states, whereby they can co-ordinate attention between themselves, a social partner and an aspect of the environment, a state which requires them to develop communicative gestures in order to signal to that social partner what they wish to share (Cameron-Faulkner, 2020). Cameron-Faulkner et al. (2015) found that infants' use of hold out and gives was the first evidence of communicative gestures of this kind, and such gestures were a precursor to index finger pointing in that their use at 10 months were positively related to pointing at 12 months. Hold out and gives were argued to be an important precursor to pointing in that they served as 'proximal practice' for the development of the later skill (Cameron-Faulkner et al., 2015, p.584).

A longitudinal study of infants aged 11 to 41 months by Cochet and Byrne (2016) confirmed the importance of declarative pointing as an important precursor for language development, in that declarative expressive pointing was significantly associated with language production, and declarative informative pointing was associated with both production and comprehension. The study also found significant relationships between language production and the use of both head nodding and symbolic gestures, with effects greater where the infants themselves had initiated the episode of joint attention (the latter point supported by Liszkowski & Tomasello, 2011).

Word Learning

A word can be defined as the smallest linguistic unit which can be articulated in isolation to convey semantic content (Stanford Encylopedia of Philosophy, n.d.). Children's ability to understand and speak whole words is integral to their language development and overall ability to communicate, alongside their understanding of grammar and pragmatics.

Receptive and Expressive Language. Whilst language develops holistically, with some authors claiming that language skills overall develop as a 'unitary construct' (McDonald, 2013, as cited in Dimova et al., 2020, p.16), with component parts often difficult to extricate from one another in practice, it is nonetheless important to consider the distinction between receptive and expressive language. Put simply, expressive language refers to the words children can say – the production of language (Harley, 2001, p.420), and receptive language to the words children understand - the comprehension of language (Harley, 2001, p.423). Both receptive and expressive language development are integral in early language acquisition, including the development of pre-lexical skills (for example babbling and recognising sound patterns), although their influence may vary depending on the particular stage and skill being developed. Using CDI data for a sample of 1803 children aged between 8 and 30 months old, Bates et al. (1995) found that, for most children, evidence of word comprehension appears between 8-10 months old, and production at 12 months old. It has been suggested that the development of receptive and expressive language skills may draw on different cognitive resources, and be mediated by different neural systems, with non-linguistic measures (such as communicative and symbolic gestures) tending to correlate more strongly with the understanding of language, as opposed to its expression (Bates et al., 1995). The implications of this have not been fully explored, but the relationship between the two, and how each are measured, should be borne in mind as part of any investigation into language.

As noted above by Bates et al. (1995), it is generally accepted that receptive vocabulary tends to precede expressive vocabulary, and be larger in size during the early stages of language acquisition (Clarke & Hecht, 1983; Gershkoff-Stowe & Hahn, 2013). The former assertion is likely due to the fact that speech production tends to require the speaker to have some understanding of the word in question in order to communicate effectively (Clark, 1995 as cited in Ryan et al., 2016) – although this is not always the case (for example a child may imitate a word or phrase and have no understanding of its meaning). Further insight is offered by Sénéchal (1997) in the context of investigating the effect of a shared reading intervention on both receptive and expressive vocabulary. For receptive vocabulary knowledge, children were required to encode (requiring a comparison between external and internal representation), store and retrieve information, something which repeated exposure to

books was likely to facilitate. Expressive language, on the other hand, required the additional process of reproducing a phonological representation of the encoded word, something which was more likely to be facilitated by opportunities to interact and practise such a skill. Despite the general acceptance that receptive vocabularies come first and are larger in size, there are some exceptions to this, and in some contexts production skills may exceed comprehension skills – for example specific areas of syntactic development (Matthews et al., 2009), for autistic children (Hudry et al., 2010) and for children from lower SES backgrounds (Ryan et al., 2016) - suggesting a somewhat more nuanced relationship between the two. Nonetheless, some evidence suggests that delays in early language comprehension may be a more reliable indicator of language delay than production alone (Rescorla, 2009), although, as noted by Frank et al., (2021), receptive language scores via parent report may be more susceptible to bias, in that that 'comprehension' might mean different things to different parents, whereas reported speech carries more signals and is therefore more straight-forward.

Laying the aforementioned disassociations between receptive and expressive vocabulary aside for a moment, it is logical to assume that speaking and understanding do not develop separately (Bloom, 1974), supported by evidence that measures of receptive and expressive vocabulary tend to be highly correlated (Frank et al., 2021). Indeed, it has been suggested that the earlier development of receptive language may even aid the development of expressive language skills (Gibson et al., 2012), although the mechanism through which this happens is unclear, and the words children produce are not necessarily a subset of their receptive vocabularies (Goldin-Meadow et al., 1976). The link between receptive and expressive vocabulary is not routinely considered within the existing literature. For example, studies which have established a link between child-directed speech (CDS) and expressive vocabulary (e.g., Hoff & Tian, 2005) do not regularly consider the mechanism through which children gain receptive vocabulary knowledge via the verbal input they receive, and then convert this receptive knowledge into expressive language.

One possible explanation for the link between receptive and expressive vocabulary is related to processing efficiency, which in the context of children's language acquisition refers to the speed in which they can process incoming information and demonstrate their understanding, and is measured by both speed and accuracy (Fernald et al., 2013). A child's ability to efficiently process verbal information from their environment is closely related to the development of their receptive vocabulary, and Marchman et al. (2016) found lexical processing speed to be the strongest predictor of receptive vocabulary in a sample of 30 preterm 36-month-olds. The more verbal input children receive, the more opportunities they will have to practise the necessary skills, including segmenting speech words and accessing lexical representations – allowing them to 'orient to familiar words more quickly and accurately' and thus 'facilitate rapid vocabulary growth' (Weisleder & Fernald, 2013, p.2149). But how does this improvement in receptive vocabulary, achieved via a combination of verbal input from their environment and ever-improving processing efficiency, contribute to language production? Fernald et al. (2006) were interested in this very question, and specifically how changes in receptive vocabulary between 12 and 24 months related to rapid development in expressive vocabulary over the same time period. In this study, a significant relationship was found between speech-processing efficiency and vocabulary production for 21-month-olds and, crucially, was explained via the wider benefits of processing efficiency, in that general improvement in working memory and cognitive competencies brought about by efficient processing of verbal input (Fry & Hale, 1996; Bates et al., 1988) could facilitate a range of competencies within language and beyond it

(including expressive vocabulary). Additionally, increased opportunities for hearing, learning and imitating were likely to create positive feedback loops, with greater understanding eliciting more CDS and more complex parent-child interactions, thus leading to 'cascading advantages' for children's receptive and expressive vocabulary (Fernald et al., 2006, p.114).

To summarise, receptive and expressive language appear to be both linked and disassociated to some degree. The mechanism through which they are linked is under-defined, but could be something to do with processing efficiency, in that the ability to process language input quickly and accurately not only aids receptive word knowledge, but also drives expressive vocabulary via general cognitive improvements. Cascading advantages could ensue for those children who show greater levels of understanding and production, eliciting more talk from their caregivers, and more opportunities for interaction and imitation, leading to even better processing efficiency. The link between receptive and expressive language is an important consideration for constructivist accounts of language which place great importance on the language environment.

Different Theoretical Approaches to Word Learning. The major discrepancies concerning the role of nature and nurture within general language acquisition tend to focus on the acquisition of grammar rather than vocabulary, and it is the latter of these that is of interest here. In terms of vocabulary development, there are three main competing explanations: the lexical constraints model, social-pragmatics and associative learning accounts. These accounts primarily refer to how children acquire receptive vocabulary knowledge, as opposed to how they perform expressively (with the inherent links and differences between the two discussed in the section above). The lexical constraints or principles model states that children are able to acquire words by utilising numerous default assumptions (that may be innate or learned), including the whole-object assumption, meaning that a word will tend to refer to the whole object, not a part of it (Golinkoff et al., 1994), the taxonomic assumption, meaning that a word can refer to other objects of the same kind, even if they do not appear the same (Markman, 1989) and mutual exclusivity, whereby a new word will refer to a different object (Merriman & Bowman, 1989). One criticism of this approach is that children regularly do not abide by these given constraints, and are, for example, well able to learn parts of objects, and multiple names for the same object, although proponents assert that it is possible to override constraints where counterevidence exists (Woodward & Markman, 1991).

Social pragmatics is an alternative approach for explaining word learning, whereby children employ skills of joint attention and understanding communicative intention to acquire word meanings (Akhtar & Tomasello, 1996; Baldwin et al., 1996). This approach removes some of the ambiguity involved in learning words: children are no longer forced to decide between an infinite number of possible referents for the words they hear - the so-called Quinean conundrum (Quine, 1960) – as the adult's communicative intentions make the word's referent clear (Bloom, 1993). The fact that these socially pragmatic skills and first words begin to emerge at the same time, at around 12 months old, is not merely coincidental, and provides further supporting evidence for the approach (Tomasello & Haberl, 2003). Conversely, the fact that other studies that have found children able to learn words before they acquire the socially pragmatic skills central to this theoretical approach appear to contravene it (Floor & Akhtar, 2006), in that such skills may be useful, but not sufficient for word learning at all developmental stages, suggesting the need for a more finely-grained explanation, rather than a one-size-fits-all, at every age, approach.

Finally, the associative learning approach (or cross-situational statistical learning) is whereby children acquire word meanings through the observations of the co-occurrence of words (or by their absence) and an example of their meaning. Whether or not a purely associative approach is capable of explaining all word learning – in particular verbs or more abstract nouns – is debatable – indeed the associative learning approach is often utilised as an explanation in conjunction with syntactic bootstrapping (Gleitman, 1990), whereby learners use the syntax and grammar of the sentence to 'bootstrap' the word's meaning, particularly verbs (Naigles, 1990), although it should be noted that in principle, bootstrapping could work in conjunction with any one of the three mentioned theories.

Various debates have ensued regarding which model or combination of models provide the most plausible explanation for vocabulary learning. One example of the latter is that the 'constraints' in the lexical constraints model are more likely to be general learning principles either socio-pragmatic or associative in nature – or possibly both – with some kind of bootstrapping used from around 2 years of age (Ambridge & Lieven, 2011) - another is the idea that all learning is in fact associative, with social pragmatic learning simply a higher level description of associative phenomena, (Landauer & Dumais, 1997). The Emergentist Coalition Model (ECM) of word learning (Hollich et al., 2000) is a hybrid model drawing together some of the competing principles outlined above, that emerged from a 'growing consensus' (Golinkoff & Hirsch-Pasek, 2007, p.1) that a single approach was not sufficient to explain children's lexical acquisition. Thus, the ECM asserts that word acquisition is a product of cognitive, social and linguistic cues, rather than one or the other, and the relative influence of such cues will differ depending on the child's developmental stage. Younger learners are more likely to rely on attentional cues such as perceptual salience, before becoming more reliant on linguistic and social cues, (Hollich et al., 2000, p.24), although in practice different cues may also be weighted from moment to moment, not just at different developmental stages (Ambridge & Lieven, 2011), and parallels can be drawn here with the multiple cues used when segmenting words from the speech stream (Romberg & Saffran, 2010). Importantly, these cues (or 'lexical principles') exist only in an immature way at the outset, and thus need to 'emerge' from the word learning experience, thereby supporting both nativist and constructivist positions (Golinkoff et al., 2000). The ECM examines language from the perspective of comprehension rather than production, using an Interactive Intermodal Preferential Looking Paradigm (IPLP) to test its parameters. Having said this, the ECM has been used as a theoretical framework in studies looking at both receptive and expressive vocabulary (see e.g., Maekawa & Storkel, 2006). The receptive vocabulary knowledge gained via use of multiple cues (as per the ECM) could then presumably be translated into expressive vocabulary via imitation and practise, as noted above in the discussion of the relationship between receptive and expressive language.

Fast Mapping. Models of word learning with constructivist elements, including the ECM, assume repeated exposure to words is necessary for word learning, although how many repetitions are necessary is disputed, and the quality of the exchange in which a child hears a word is also important alongside how many times they hear it (Hindman et al., 2016). The notion of repeated exposure being necessary for learning has, however, been challenged by the suggestion of a 'fast mapping' procedure, first noticed by Carey and Bartlett (1978), whereby some children were able to learn and retain novel words after only one exposure in experimental conditions. This idea was supported by Spiegel and Halberda (2010), whose experiment showed that 2 year olds were able to fast map novel nouns after a brief single exposure, and by Cartmill et al. (2013), whose investigation into the importance of the

quality of the exchange suggested that a single high quality word exposure, supported by non-verbal cues and referential transparency, was preferred to many lower quality exposures, thereby supporting a fast mapping procedure. Fast-mapping may be more suited to laboratory settings however, where word knowledge tends to be tested in the short-term, with Dockrell et al. (2007) demonstrating that fast-mapping procedures were not sufficient for key stage one pupils attempting to learn new scientific vocabulary from an educational video, possibly because such terms differed from the conceptually simple words learned in early childhood, often object names directly experienced in the immediate environment. Further doubt has been cast on the effectiveness of the fast mapping process by the finding that words learned in this way may also be forgotten more quickly (Vlach & Sandhofer, 2012). Nonetheless, here the authors asserted that the process of forgetting is not necessarily a contravention of successful word acquisition, and actually facilitated extended mapping (whereby children refine their knowledge of a word through many repeated exposures) and generalisation of words and categories. The evidence therefore appears to suggest that fast mapping (or even just fast*er*, rather than fast) may occur in certain circumstances, for example when words are conceptually simple (Dockrell et al., 2007) or are particularly well-embedded within a high quality and useful context (Cartmill et al., 2016) – suggesting that frequency of input is not the only important factor when considering how quickly and easily children acquire words.

What is a Typical Vocabulary Trajectory? There are numerous sources which signpost a broad range of language milestones which a child is expected to reach between the ages of 0-5 years, including Great Ormond Street Hospital (GOSH), Children's Communication Charity: I CAN (ICAN), and the American Speech-Language-Hearing Association (ASHA). Whilst these milestones are broadly consistent, for example regarding the development of non-verbal abilities, when first words begin to emerge (at around the 1-year mark) and at what stage children are able combine words into short utterances (between 18-24 months), they are less consistent regarding the number of words a child should know and be able to say at each stage. Indeed, these sources often avoid stating specific numbers for receptive and expressive vocabulary sizes altogether, suggesting either uncertainty and/or a wide range of what is considered typical at each stage. For example, ASHA's guidelines suggest that by 24 months children should be able to use and understand 50 different words, whereas ICAN's guidelines suggest that children should be able to produce 50 words, but understand far more than that by 24 months – around 200, and 300 by 36 months. Fenson et al. (1994) used parental report data (using the MacArthur-Bates Communicative Development Inventory) from 1803 8- to 30-month-olds, and concluded that there was a high level of variability in children's lexical development within this age range: productive vocabulary for a 24 month old ranged from under 100 words to over 500, with similarly wide ranges for 16 month olds and 30 month olds. The question of a 'typical' vocabulary size, either in productive or receptive vocabulary, thus remains somewhat elusive.

Different Approaches to Measuring Vocabulary. There are many different methods and contexts in which language can be measured, either globally or within specific domains, including expressive and receptive vocabulary. Standardised instruments designed to measure the developing vocabulary of a child tend to focus on either receptive vocabulary or productive vocabulary. Examples of measures for receptive vocabulary include the Peabody Picture Vocabulary Test (PPVT-5, Dunn & Dunn, 2007) and the British Picture Vocabulary Scale (BPVS, Dunn et al., 2009), the former of which is US-normed and the latter UK-normed. Examples of measures for expressive vocabulary include the British Ability Scales II Naming Vocabulary (BAS_II, Elliott, 1996), and the Renfrew Word Finding Vocabulary Test

(RWFVT, Renfrew, 1995), both of which are based on UK norms. These standardised tests tend to be administered by professionals, although they can still be susceptible to bias, for example tests of receptive language which tend to involve listening and then pointing at pictures or objects may lend themselves to a guessing strategy, which wouldn't work for equivalent tests of expressive vocabulary, where words need to be generated from scratch (Leonard, 2009, as cited in Gibson et al., 2012).

Whilst also a standardised measure, the MacArthur-Bates Communicative Development Inventories (Fenson et al., 2007) rely on parent report to measure both receptive and productive vocabulary across a range of languages from 8 months old (the other measures mentioned here are designed for older children, although intended age ranges differ between measures). The benefits and drawbacks of parent report measures are further discussed in Chapters 3 and 6, but it is worth noting again here the different way in which parents may report production and comprehension as discussed by Frank et al., (2021) may mean that receptive language scores are more susceptible to over and under-estimation due to the fact that 'comprehension' is harder to ascertain, and might mean different things to different parents.

Measures of vocabulary based on standardised tests necessarily act as indexes of vocabulary, and ceiling effects will be apparent as children approach the upper end of the intended age range, or for those with particularly large lexicons to begin with. Academic research has adopted other formats of measuring vocabulary, including naturalistic observational studies and experimental paradigms. One example of the latter is the Interactive IPLP paradigm used to test receptive vocabulary via the ECM framework (Hollich et al., 2000), and using naturalistic language samples to study expressive language will be discussed in more detail in Chapter 6. Beyond that, there is a wider discussion to be had around the pros and cons of each method and context for measuring vocabulary that is beyond the scope of this thesis, nonetheless it is important to note that children's language scores tend to be positively correlated regardless of the measurement tool or method used (Bornstein et al., 1998). In the latter study, six measures of vocabulary taken from three different sources (spontaneous speech, experimenter assessments and maternal reports) for a sample of 131 children aged 1 year and 8 months all co-varied with each other.

Finally, further consideration should be given to the distinction between measures of expressive language, receptive language and depth of knowledge (a more highly developed understanding of how a word is used in different contexts which could be applied to both expressive and receptive language). For example, is a child able to use a particular word correctly in multiple contexts? Can a child apply their understanding of a word to encompass many different types within that same category? If a child produces a word, or can be shown to understand its meaning on one level or other, then Ambridge and Lieven (2011) argue that this marks merely the beginning of the acquisition process, through which, over a period of time, the meaning of the word continues to be refined and the subtleties of how the word can be used in different contexts learned (Harley et al., 2018). There are very few studies that measure depth of knowledge as an outcome measure of vocabulary (e.g., Agdam & Sadeghi, 2014).

Development of Grammar: Syntax

Sentences are not just a collection of words, and whether they are simple utterances or more complex forms, they contain additional grammatical information which communicates further

information about, for example, whether the sentence is a question or a statement, who did what to whom, and when something took place (Ambridge & Lieven, 2011). Such grammatical information can be communicated both via morphology (for example the use of suffixes to indicate verb tense) and syntax (for example knowledge about syntactic categories, and the usual order of these categories within a simple sentence), with the latter being of overriding importance in the English language. Children must acquire knowledge of these grammatical rules in order to communicate effectively.

In terms of the acquisition of early syntax, one of the key considerations is how children begin to understand syntactic categories (for example noun, verb, adjective, etc), and how they assign each new word they learn to one of these categories, and thus are able to use it effectively (in the correct position) within a sentence. The default construction for simple sentences in the English language is subject-verb-object (SVO) (e.g., *the girl kicked the ball*), although this default construction becomes more nuanced as sentences become longer and more complex, and it is not necessarily the common construction of CDS utterances (e.g., Cameron-Faulkner et al., 2003, found only 15% of maternal utterances to follow the SVO format). Theories on how children acquire syntax are more traditionally divided along nativist-constructivist lines, for example Pinker (1984, 1989, as cited in Harley, 2013) argued for semantic bootstrapping, which relied on children having innate knowledge of syntactic categories and suggested that they were able to make an inference about the underlying structure of a sentence based on its surface structure and knowledge about its meaning. An alternative explanation is that no innate principles are required, and that knowledge of syntactic categories are constructed based on semantic knowledge - for example many nouns describe concrete objects, and many verbs describe actions - although the fact that this is not always the case is a limitation. A third explanation to be considered is distributional analysis, whereby children are thought to statistically analyse the input they receive looking for patterns which will indicate a word's syntactic function. As noted above, CDS does not always follow a consistent SVO format to allow easy extraction of patterns, but input can also come from other sources which are more consistent in this respect (e.g., shared reading, as suggested in Cameron-Faulkner & Noble, 2013), and the information extracted can also be morphological in nature or related to stress and syllable patterns – for example common suffixes *-ing* and *-ed* tend to indicate verbs (Maratsos, 1982, as cited in Harley, 2001), and nouns and verbs tend to have different stress patterns (Kelly, 1992) – thus allowing children to learn these categories regardless of semantic information or innate learning principles. What is also clear is that, regardless of which process is responsible for facilitating children's acquisition of syntactic categories, as well as their subsequent use in helping them to understand and produce simple sentences, this process is difficult to extricate from other domains of language development, including lexical development, where clear relationships between grammar and vocabulary development have been found (Braginsky et al., 2015).

Pragmatic Skills

Pragmatic skills are concerned with a person's ability to use language effectively to communicate across different situations, including asking and responding to questions, taking turns in a conversation, telling jokes and using an appropriate volume and tone of voice (Bryant, 2009). The initial stages of pragmatic development begin in the first year of life with the development of non-verbal communication skills (as discussed above), which mark the onset of triadic joint attention states and both understanding others' and indicating personal communicative intentions. This can be

demonstrated, for example, by a toddlers' ability to persist with a conversation topic even after a non-acknowledgement during an interaction with their mothers, although they were less persistent with their fathers, who they had worked out were more likely to follow up their attempts with a further non-acknowledgement (Tomasello et al, 1990).

The ability to initiate, sustain and take turns in conversations is a particular focal point for researchers interested in pragmatic development. Pre-schoolers rarely overlap when taking part in conversations with each other (Bryant, 2009), although pauses between turns can be particularly long (Garvey & Berninger, 1981, as cited in Hilbrink et al., 2015), perhaps because children wait until their partner has stopped speaking before formulating a response. Hilbrink et al. (2015) found that turn-taking actually slows down (involving longer pauses) before infants produce their first words, before speeding up again later on – with later research suggesting that such pauses were in fact optimally timed, as they allowed time for pre-schoolers to coordinate language comprehension and production (which necessarily becomes more complex as they get older) and prepare a response whilst still listening (Lindsay et al., 2019). Toddlers and pre-schoolers will often use repetition and recasting of partner utterances to keep a conversation going, (Pan & Snow, 1999), but as they get older they are able to add new information to each turn. Indeed, pragmatic development continues well beyond childhood and into adolescence and beyond, with teenagers developing their conversational abilities by providing more feedback (such as nods and comments) and using more sophisticated connectives to keep the conversation going (Hoyle & Adger, 1998), as well as developing competence in the use of slang terms and being able to communicate effectively across different modes, notably those afforded by newer technology.

As with other domains of language, pragmatic skills are not always easily extricated from other aspects of language development, although it is possible for children to develop skills in some aspects of language and not master pragmatics (e.g., Blank et al., 1979). Nonetheless, all aspects, including pragmatic ones, are clearly required for a comprehensive mastery of language and communication, and links between vocabulary acquisition and pragmatic development, for example, are clear in that children tend to acquire lexical items of particular pragmatic importance (such as *hi*, *bye* and *thank you*) relatively early on (Fenson et al, 1994), as well as words that are embedded in high quality interactions (Weizman & Snow, 2001). What is also worth noting here is the importance of context and wider culture in developing those pragmatic skills, and how this interacts with other aspects of language development as noted. Pragmatic behaviours are 'contextually sensitive' (Bryant, 2009, p.341), and thus difficult to plot on a developmental timescale – for example a child who is able to take appropriate conversational turns may only do this with familiar adults, and learning to use the words *please* and *hello* are fraught with difficulties as their use depends on the nuance of the individual context, as well as how these pragmatic skills have been taught. In Becker (1994), parents were observed to comment on wide variety of pragmatic behaviours (for example prompting the use of *please* and *thank you*, or commenting on volume and tone), suggesting that, as with other aspects of language, primary caregivers are of pivotal importance concerning when and how children develop skills which are important for the pragmatic use of language, and how and when they use such skills will reflect the culture and context in which they were learned.

Summary of Chapter 1

Early language difficulties have been linked to numerous short and long term challenges for children both in school and beyond, and thus helping children to overcome such difficulties has become an important political priority. Language difficulties, or SLCN, can encompass a broad spectrum of needs. Children with smaller vocabularies than their peers, or 'language delay,' can also be considered to have a SLCN, even if there appears to be no specific reason for this delay beyond the environmental input they are receiving. These children are particularly relevant when studying the vocabulary gap, which will be considered in detail in Chapter 3. Despite children's language difficulties appearing to be widespread, there are challenges in identifying and targeting the children most in need of support before they reach school. This is partly due to there being only one official check of oral language skills before children start school, and yet there is potential for an increasingly important role for early years practitioners in terms of identifying and supporting language difficulties, particularly as the effects of 30 hours free childcare come into play.

Theoretical approaches to how children acquire language (words and grammar) have traditionally split along the nature/nurture divide, although a complex interplay is the likely explanation. For word learning specifically, a number of theoretical approaches have been presented, including the Emergentist Coalition Model (ECM) of word learning, which posits that children learn word using a combination of cognitive, social and linguistic cues, with each asserting varying influence depending on both the context of the interaction and the developmental stage of the child. Learning words is just one of the stages of children's language acquisition, which begins as soon as children are born, if not before. Skills such as babbling, phonological development, segmenting from the speech stream, non-verbal abilities (such as the ability to engage in triadic joint attention states), syntactic abilities and pragmatic skills develop before, after or alongside the ability to speak and understand words. Consideration was also given to the distinction and relationship between receptive and expressive vocabulary, both in general, and with respect to the different approaches to measuring vocabulary.

The next chapter will consider the importance of a child's environment in the context of language learning, in terms of the quantity and quality of verbal input they receive, and which environmental factors may be associated with language outcomes.

Chapter 2: The Importance of a Child's Verbal Environment

Children from across the globe have been found to meet developmental milestones, including those related to language, at similar ages (Chomsky, 1957), and yet it is also clear that there is variation in the rate and possibly trajectory at which children acquire vocabulary (Fenson et al., 1994). This difference could be minor and self-correcting, as appears to be the case with late talkers who subsequently catch up with their peers (Girolametto et al., 2001), or pervasive and long-lasting, as is often the case with developmental language disorders (Bishop, 2017) and those with language delay due to differences in environmental input (Finnegan & Warren, 2015). As we have seen, theories of language acquisition place different emphases on genetic and environmental influence, although with vocabulary development (as opposed to grammar), there is more consensus over the importance of environmental input, as word labels and meanings cannot exist innately (Ambridge & Lieven, 2011), and depend upon the language being spoken. Focusing on a child's language environment (and any circumstances which may influence this) is therefore consistent with the overarching aim of this thesis and any follow up research: namely, to benefit children's vocabulary and overall language outcomes. There is potential to achieve this aim by focusing on a child's language environment (as opposed to genetic or biological influences), particularly the verbal input a child receives, and how specific environmental circumstances, such as poverty, the number of children in a household or access to libraries (for example), may affect this input. This chapter will consider research which has linked both quantity and quality of verbal input to children's language outcomes, as well as the possible mechanisms explaining this link. It will then go on to consider a range of demographic factors which have been associated with children's language. The terms 'parent' and 'caregiver' will be used interchangeably throughout this chapter, and in instances where one has been used and not the other, then this has been done in the pursuit of conciseness, and in most cases, both are intended.

Research Linking Quantity of Verbal Input to Child Language Outcomes

Up until the late 20th century, it was assumed that individual differences in vocabulary depended mostly on variations in an individual's cognitive learning capacity, yet in the first published study of its kind, Huttenlocher et al. (1991) found that there was also a substantial relationship between the quantity of maternal talk a child was exposed to, and the size of the child's own vocabulary. More weight was given to this claim by Hart and Risley (1995), who found that there was a huge variation in the number of words 3-year-old children heard per hour, ranging from 620 to 2150, and that these numbers appeared to be correlated with the socio-economic status (SES) of the family, with lower income families tending to speak less to their children. Importantly, the number of words spoken to the child also correlated with the size and diversity (number of different words) of the child's own subsequent vocabulary. The much quoted '30-million-word gap' (in terms of word tokens heard by children) was subsequently extrapolated from these findings, implying that children from lower income backgrounds heard 30 million less words than those from higher income backgrounds. The Hart and Risley study has been cited over 8000 times (according to Google Scholar) and their findings both robustly supported (e.g., Hoff & Tian, 2005; Huttenlocher et al., 2010; Fernald et al., 2013) – yet also questioned. Indeed, research in this area has also found significant within-group variation when looking at homogenous SES groups, suggesting a more complex relationship between social groups and language outcomes, with many other factors independent of SES having the potential to affect

child-directed speech (CDS), including parental mental health and conversational style (Weisleder & Fernald, 2013).

Criticism of the Hart and Risley study tends to focus on either its methodology, the interpretation of its results, or the deficit perspective some argue it promotes. Sperry et al. (2019), for example, failed to replicate the Hart and Risley results, and argued that a child's language environment could only be accurately measured if overheard speech was accounted for alongside CDS, as both elements of verbal input played an important role in language development. The importance of overheard speech is demonstrated by Mayan children, who reach similar developmental milestones at appropriate ages, but are rarely addressed directly by adults (Schneidman & Goldin-Meadow; 2012, Brown & Gaskins, 2014). Sperry et al.'s position on overheard speech was contested by Golinkoff et al. (2019), who argued that numerous studies had evidenced that it was the interactive nature of contingent back-and-forth conversation, rather than input alone, which fuelled language learning (Goldin-Meadow et al., 2014; Tamis-Lemonda et al., 2014). Overheard speech required a higher attentional and socio-cognitive demand from a child (including having to infer another's interests and intentions from afar) and was missing the characteristic properties of CDS in terms of simplicity of content and grammar, prosody and physical cues that have been linked to child language outcomes (Golinkoff et al., 2015). Whilst several experimental studies have shown some success in young children learning from overheard speech (Messenger et al., 2015), these tended to control or minimise attentional distractions, thereby reducing their ecological validity in a more naturalistic setting. On balance, therefore, whilst indeed it may be possible for some language learning to occur via overheard speech, the majority of the evidence suggests that children benefit most from speech that is intended for them.

Some critics of Hart and Risley have also found issue with the extrapolated figure of 30 million, which unrealistically assumed a consistent number of words heard across each day, every day (Purpura, 2019). This observation was supported by D'Apice et al. (2019), who conducted naturalistic observations using LENA technology of 107 families over 3 days for 15 hours each day (the largest to date) and found that home language input varied as much within the same family on different days as it did between different families, with certain daily activities yielding more adult speech than others. Indeed, the finding that language input was, perhaps rather obviously, a dynamic process that evolved over time and was dependent on the context, made extrapolating accurate findings from 1-hour recordings rather difficult. In Hart and Risley, the participant families were recorded for 1-hour sessions, and, assuming consistency across hours of the day, results were then extrapolated to provide a guide for daily input. These hourly sessions were conducted at a time convenient to the participant family, and involved more intrusive recording technology than is available today – it is therefore perfectly conceivable that there was a degree of Hawthorne effect involved for some families, thereby over-estimating the level of CDS for the whole day whole day, as well as potentially exacerbating any differences found. Yet exacerbating the magnitude does not negate the existence of an input gap altogether, although this is clearly only part of the story in explaining the vocabulary gap, as will be discussed in subsequent sections regarding quality of input.

One final theme of reproach directed at the original Hart and Risley study (and indeed others that have followed its lead) is the language used to describe the phenomenon of varied language input and output. Many researchers have taken issue with the widespread use of the terms *gap* and *deficit*,

claiming that they alienate the very communities that they are designed to serve (Anderson & Linden, 2016), and do not appreciate the diversity of different cultures and language practises that go beyond vocabulary (Cartmill, 2016). This is an important point to consider (and will be revisited in Chapter 10), as whilst the use of such terms may seem relatively innocuous in academic circles, once, as is apparent here, they become part of the public rhetoric, via politicians, media outlets, schools and charities, their use may have a more provocative effect, and will most likely be detrimental to the cause of recruiting, engaging and motivating families to adopt different behaviours. If a trend of more positive language were to emerge, coupled with an increased appreciation of different cultures, this could have a positive effect on both how researchers collect information, and how interventions are designed in culturally sensitive ways for maximum engagement and impact. Whilst reframing the issue within positive language, or even reconceptualising gaps as debts – in education, teacher quality, school funding or income, as suggested by Milner (2013), is unlikely to result in significant changes to the education and assessment system, it could, over time, lead to a shift in the onus of responsibility and an increase in funding.

Ultimately, under the current system, children are expected to reach distinct milestones at specific ages throughout their childhood, as evidenced by the use of statutory and non-statutory screening and testing within the UK education system and beforehand, including the ASQ, EYFSP, Standard Assessment Tests (SATs) testing and GCSEs. Those who do not meet these expected standards are likely to face short-term and long-term repercussions, including loss of self-esteem and a narrowing of the further education and job prospects which are open to them. The myriad criticisms directed at the original Hart and Risley study raise some interesting points about how the word gap is conceptualised and described, and challenge its magnitude, yet they do not challenge the existence of a word gap in terms of a wide variation in words spoken by caregivers to children, and indeed later studies have confirmed such a an input gap to exist, sometimes associated with SES (Hoff 2003; Huttenlocher et al., 2010; Huang et al., 2017; Gilkerson et al., 2017), although not always (Weisleder & Fernald, 2013) and this input gap appears to have a subsequent effect on children's own language – hence the resulting vocabulary gap.

Research Linking Quality of Verbal Input to Child Language Outcomes

Subsequent interactionist research, whereby differences in input are associated with differences in output, moved away from focusing on the quantity of input and towards a focus on quality, with 'quality' referring to the specific words used, how they were used, plus the non-verbal features of the communication exchange.

Diversity of Input

Weizman and Snow (2001) found that children's vocabulary scores were more substantially related to maternal use of sophisticated words than to the quantity of words spoken to them overall, with 'sophisticated' referring to any word that fell outside the 3000 most common words contained within the Dale-Chall wordlist. *Vehicle*, *tusk*, *cholesterol* and *gulping* are given as examples of such 'sophisticated' words (Weizman & Snow, 2001, p.269). Not only that, but it was the extent to which such words were embedded in helpful and instructive interactions that predicted vocabulary two years later. Huttenlocher et al. (2010) corroborated the link between diversity of earlier caregiver speech and corresponding diversity in later child speech, with earlier child speech also predicting later

caregiver speech, suggesting a multidirectional influence and a spiral of accumulating advantages for children who are the beneficiaries of diverse caregiver speech in their first years.

In an attempt to separate the effects of quantity and quality of input on language learning, Rowe (2012) measured input in terms of rare words and decontextualised language, whilst controlling for SES, quantity and previous vocabulary knowledge. Results suggested that the relative importance of input quantity and quality was dependent on a child's developmental stage, with quantity and repetition more productive during the 2nd year, variation or 'quality' in the 3rd year, and narrative in the 4th year. These results corresponded with those of Jones and Rowland (2017), who provided further evidence for the claim that lexical diversity in input (the variety of words spoken to a child) was significantly associated with lexical diversity in output (a child's vocabulary scores) via a computational model designed to mimic child language acquisition. Whilst quantity of input was again deemed to be important in the early stage of learning, lexical diversity was ultimately more important as it allowed the model to accumulate a larger store of sub-lexical chunks that could then be utilised to process new input – again leading to a 'virtuous circle' (Jones & Rowland, 2017, p. 19) whereby the more lexical diversity a child was exposed to, the more sub-lexical information they could call upon, the quicker their processing speed and the more efficient their vocabulary learning.

In summary then, lexical diversity is clearly important, but it is does not provide a full explanation. Averaged across the relevant literature, Jones & Rowland (2017) suggested that lexical diversity explained only 20% of vocabulary learning, with additional variance likely to be explained by the child's prior knowledge, as well as other cognitive, environmental and genetic factors – all of which should be considered for a full model of vocabulary learning (Jones & Rowland, 2017, p.18). This is relevant not only for the sake of an accurate model of language learning in itself, but also worth considering when predicting the impact of interventions: a high quality, theoretically based intervention may not have the desired significant impact if other factors are at play and have not been, or cannot be, taken into account.

Quality of the Communication Exchange

It is not only the diversity of the words used that encapsulates a high-quality exchange. Cartmill et al. (2013) found that the variability in a child's lexicon on entering school was also explained by differences in referential transparency, meaning how clearly a word meaning can be inferred from the non-verbal features of the communication exchange. Here, in order to avoid using a detailed and ranked coding system of non-verbal cues, the authors used a Human Simulation Paradigm, whereby they measured the time taken for an adult to correctly infer a word meaning using only non-linguistic cues. Again, the importance of learning context was alluded to, with results suggesting that a single high quality exposure, embedded in a relevant context with superior referential transparency, was more productive than multiple lower quality exposures, a point corroborated by Weizman and Snow (2001), whose study showed that contexts where interactions invoked some kind of physical or social context, prior knowledge or semantically related information were more likely to facilitate learning of unfamiliar words than sheer quantity of repetitions.

Besides referential transparency, other features of the quality of the exchange have also been investigated. Schwab and Lew-Williams (2016) found that the typical features of CDS, characterised by particular prosodic features (higher pitch, exaggerated vowels, final word lengthening, louder

volume), shorter utterances, simple sentence frames and more repetition, contingency and transparency, all had a significant effect on later vocabulary development. Diversity of input in terms of syntax, as well as words, has also been found to be predictive of later diversity of child speech (Huttenlocher et al., 2010), with both constituent and clausal diversity evaluated.

Contingency, referring to talk that focuses on what a child is interested in at that very moment, also appears to be a relevant factor in children's language learning. McGillion et al. (2017) found that, for the first year of life, children from lower SES backgrounds hear less contingent talk, but that an intervention aimed to improve such talk in this group was effective at facilitating vocabulary growth, but only in the short term. Various mechanisms for why contingent talk may be particularly beneficial for vocabulary learning were suggested by Tamis-LeMonda et al. (2014), and included a proposal that the temporal characteristics of contingency increased the likelihood that infants will link the words they hear with their correct referents in the real world. Furthermore, the evidence for contingency provides further support for the need for a rich and varied learning context: if children learn best from their immediate surroundings (Tamis LeMonda et al., 2014), then it follows that word learning is most effective when it takes place in an analogous physical context, for example learning the word *zebra* at the zoo (Cartmill et al., 2013). Providing such rich and varied contexts is of course much more difficult for those without the financial or educational means to do so, although not impossible.

Common sense dictates that for contingent to be effective, there must be some element of joint attention. Joint attention has been independently linked with both language comprehension and production (Cochet & Bryne, 2016), although only when it had been initiated by the infant, a skill that implied an understanding of others' intentions and motivations, and possibly facilitated by imitation skills that further 'promoted affiliation with others' (Cochet & Bryne, 2016, p.195). These results, suggest that the success of contingent talk as a facilitator of language learning relies on infants developing the prerequisite skill of establishing joint attention beforehand.

Possible Mechanisms Explaining the Link between Verbal Input and Language Outcomes

The large body of evidence that links both the quantity and quality of input (with quality including lexical diversity) to a child's subsequent language outcomes does not contradict any genetic or biological influence on language outcomes, but merely confirms the important role that the environment plays. As noted in Chapter 1 when discussing the differences and relationship between receptive and expressive language, a number of suggestions have been made regarding the mechanism through which verbal input can influence both receptive and expressive vocabulary.

One of these suggestions is processing efficiency, referring to the speed and accuracy with which children can process incoming information. Processing efficiency is positively related to both receptive (Marchman et al., 2016) and expressive vocabulary (Weisleder & Fernald, 2013). The latter study used all-day recordings of parent and child speech, finding that those children who were exposed to a higher quantity of CDS became more efficient at processing familiar words in real time, as they were provided with more opportunities to practise the skills that were vital for word learning, including segmenting speech and accessing lexical representations. This then allowed infants to orient to familiar words more quickly and accurately, thereby facilitating a more rapid growth of new vocabulary and larger expressive vocabularies at the age of 24 months. Fernald et al. (2006) also found a significant relationship between speech-processing efficiency and vocabulary production for 21-

month-olds, and explained this via the wider benefits of processing efficiency, in that general improvements in working memory and cognitive competencies were able to facilitate a range of competencies within language and beyond it (including expressive vocabulary), as well as positive feedback loops created by increased opportunities for hearing, learning and imitating, in terms of greater understanding eliciting more CDS and more complex parent-child interactions.

On the other hand, Peter et al. (2019) argued that studies showing a predictive relationship between processing speed and vocabulary development, including Weisleder and Fernald (2013), did not effectively control for concurrent vocabulary, making the resultant predictive relationship unclear, as those beginning with larger vocabularies may just have had larger vocabularies later on too. Even where studies did control for concurrent vocabulary, the relationship was not clear-cut, for example the effect was only present in those with non-typical vocabulary development, such as pre-term infants (Marchman et al., 2016) and late talkers (Fernald & Marchman, 2012). Peter et al. (2019) found a more complex relationship at play between processing speed and vocabulary development: a faster processing speed assisted vocabulary acquisition (either by allowing for more easily retrieved lexical representations or facilitating a larger store of lexical and sub-lexical chunks to utilise in new word learning), but only for children with smaller vocabularies, and other skills, such as syntactic bootstrapping, became more important once a critical mass of words had been reached. This draws parallels with the Emergentist Coalition Framework proposed by Hollich et al. (2000), whereby lexical acquisition is deemed to be the product of a variety of cognitive, social and linguistic cues which may differ in their influence depending on the child's developmental stage.

Related mechanisms have also been put forward to explain the relationship between diversity of input and child language outcomes. Jones and Rowland (2017) used a computational model to demonstrate the importance of quality over quantity as time progressed, illustrating how, once a critical level of information had been amassed through sufficient quantity of input, a more diverse input allowed for a larger number of sub-lexical sequences to be learned, thus providing a larger repertoire to be accessed and related words to be acquired more easily with fewer exposures. Those children who received a lesser initial quantity of input would, the model predicted, spend a longer period in the phase where quantity was the overriding important factor.

Summary: Quantity and Quality of Verbal Input

There is a multitude of research linking both the quantity and quality of verbal input to child language outcomes, with the latter including diversity of language used as well as the quality of the communication exchange (for example the role of prosody and contingency within CDS). In general, parents and carers who talk more to their children, include more diverse vocabulary, and embed their verbalisations in high-quality communication exchanges will facilitate positive language outcomes for their children. As also noted in Chapter 1, one explanation for this relationship is that the increased exposure provides more opportunities for interaction and imitation, leading to even better processing efficiency, which not only drives receptive vocabulary knowledge, but also expressive vocabulary via general cognitive improvements. Higher levels of verbal input could also result in a larger store of sub-lexical chunks, suggesting new words become easier to learn the larger the child's vocabulary grows. The link between a child's verbal environment and their language outcomes does not negate the probable influence of genetic factors in word learning.

Demographic Factors associated with Children's Language Outcomes

Many demographic factors have been found to be significantly associated with children's language outcomes. This could be via their effect on a child's verbal environment, or some other pathway. These factors are an important consideration when the overarching aim is to support children's language development, as they can contribute to the signposting system which directs support to those most in need. For example, Dunatchik et al. (2018) analysed regional trends in attainment using large scale datasets contained within the Millennium Cohort Study (MCS) and the National Pupil Database (NPD) and found a diverse range of factors to be significantly associated with children's vocabulary scores at 3, 5 and 7-years-old. Maternal education, employment, gender, generational status, regular breakfasts and ethnicity were all indicated as factors which influenced language outcomes and that may explain regional variations in attainment (only maternal education was consistently associated with language outcomes across region and age) – but many more factors were independently associated with vocabulary scores overall, including mother's age at birth and multiple disadvantage (a measure of SES).

A similarly wide range of variables was found to be associated with language outcomes in Law et al. (2019), this time involving data from the Avon Longitudinal Study of Parents & Children (ALSPAC), and a different measure of language (Reynell Developmental Language Scales and the CDI). Law et al. (2019) found that a combination of gender (the biggest predictor of language outcomes), prior knowledge and proximal (as opposed to distal) aspects of social disadvantage were the most influential factors, with proximal measures including noise levels, visits to the library, television hours, and perceived support available. This was a positive finding, because proximal elements of social disadvantage were more flexible to change, as opposed to more static distal elements such as income and maternal education. It is unfeasible to review the relative importance of every demographic risk factor for children's language outcomes, but using Dunatchik et al. (2018) as a starting point, a number of commonly occurring demographic risk factors will now be considered in turn.

Poverty and Socio-economic Status

There is variation in the quantity and quality of verbal input which parents and carers provide for their children (Weisleder & Fernald, 2013; Huttenlocher et al., 2010), although it is currently unclear whether such differences can be explained by specific demographic factors that align themselves with quantity and quality of input (and therefore, by default, are also aligned with child vocabulary levels). Despite this lack of clarity, poverty, income or SES are often used as a benchmark for identifying those in need, based on evidence which indicates that children from lower income backgrounds have a higher risk of poorer language outcomes (Hart & Risley, 1995; Hoff, 2003; Rowe 2017). Indeed, many charitable and political initiatives addressing the vocabulary gap and the attainment gap in general are based upon the assumption that poverty/lower SES puts children at risk of lower achievement and poorer outcomes, including Pupil Premium and various Education Endowment Foundation (EEF) intervention initiatives.

Some studies have indeed suggested that circumstances of poverty can influence the quantity and quality of verbal input parents are able to provide, for example less use of gesture (Rowe & Goldin-Meadow, 2009), and less sensitive parenting, implying less contingent talk (Leigh et al., 2011; Schwab & Lew-Williams, 2016) more directives over questioning (Leigh et al., 2011), fewer language goals in

play and more focus on directing children's behaviour (Schwab & Lew-Williams, 2016). Perkins et al. (2013) offer the Family Stress Model and Parental Investment Model as frameworks for understanding how poverty can affect parenting, with the Family Investment Model connecting the emotional stress resulting from poverty with a harsher, more authoritarian parenting style resulting in less affection and nurturing opportunities, and the Parental Investment Model illustrating how parents in poverty are forced to focus on acquiring basic necessities, and therefore lack the time and energy to provide consistent quantity and quality of verbal input.

On the other hand, authoritarian parenting styles and stressed or time-poor parents exist outside of poverty and across a range of occupations and education levels, and are not necessarily associated with adverse language outcomes. In other words, the factors that have been associated with less quantity and quality verbal input (lack of time, stress, less sensitive parenting) are most definitely not exclusive to poverty, and can exist in any parent-child relationship. The role of poverty is likely to be an exacerbating one however, in that less sensitive parenting approaches, stress or lack of time cannot be mitigated as well by a parent without money or education to compensate, and families experiencing poverty are also more likely to be affected by multiple negative influences simultaneously. For example, families with adequate resources who find themselves under stress can sometimes use money to relieve the pressure, for example by paying for quality childcare, or other goods and services that assure basic needs have been met and quality family time is ring-fenced during holidays and weekends. Similarly, those with considerable educational resources may have the knowledge to direct children to language outcomes more efficiently, or if not, know how to go about getting the necessary advice (Rowe, 2018). Indeed, the parental education component of SES has a more obvious and direct explanation regarding its relationship with language outcomes, in that the more years spent in formal education, the more extensive a person's lexicon is likely to be (which can then be imparted to their child), and possibly, the more likely they are to know about children's development. Yet knowledge and mindsets about children's intelligence and how they develop (Rowe, 2018) is not necessarily related to how many years of formal education a parent has received, and a parent's ability to incorporate stories, songs, nursery rhymes, trips out and shared activities into their daily routines, coupled with their warm and sensitive parenting style, is arguably the most important thing.

This latter suggestion appears to be ratified by studies that have focused on how the more proximal elements of social background can influence language outcomes. In a population-level study using the reception baseline assessment as a measure of school readiness and language ability, Roulstone et al. (2011) deduced that it was the children's communication environment (as measured by activities and interactions, maternal feelings, attitudes and well-being, plus resource availability) that predicted language development more strongly than the more static features of social background (as measured by parental occupation, housing tenure, overcrowding, financial difficulties and car use). These findings were corroborated by Law et al. (2019), and to some degree, Finnegan and Warren (2015). In the latter study, the authors analysed data from the National Pupil Database, finding that whilst around 40% of a 3-year-old's language skills were explained by structural factors such as income and parental education, a further 25% of the differences were explained by variables that have the potential for adaptation, such as the home learning environment and parental behaviour.

In summary, it is likely that SES, in some or all of its guises, has some influence on children's language outcomes, either via its effect on a caregiver's ability to provide consistent quantity and quality of verbal input, or via some other route, such as high noise levels or inadequate nutrition. On the other hand, the role it plays is complex, and it is likely an exacerbating factor rather than a guarantee of adverse outcomes, with proximal factors such as attitudes towards learning and daily routines being of greater importance than static factors: so what parents do is important, yet their choices can be affected by their poverty status. Understanding the role of SES in children's language development is made even more difficult by the fact that SES can be operationalised in a number of different ways, either as a single or composite measure, meaning that when studies link SES with language outcomes, it won't always be the same aspect of SES. The complexities outlined here, alongside the finding that there is as much variation in language input within SES strata as between them (Weisleder & Fernald, 2013; Hirsh-Pasek et al., 2015; Cartmill et al., 2016), suggests further research on the role of poverty/income/SES in children's language development is required.

Parental Age and Education

Parental education has been robustly linked to children's language outcomes for both mothers (Dunatchik et al., 2018; Moss & Washbrook, 2016) and fathers (Cabrera et al., 2007). In the latter study, higher levels of education indicated that fathers were more likely to engage positively with their children and provide optimal verbal input. Higher levels of parental education also meant that fathers were more likely to facilitate stimulating activities that encouraged verbal interaction (Jeong et al., 2017), as well as have a more diverse vocabulary themselves (Malin et al., 2012). As noted above however, mindset and attitude are also important contributors to a parent's ability to provide a stimulating verbal home learning environment, and incorporating stories, songs, nursery rhymes, trips out and shared activities is not necessarily dependent on years of formal education.

Maternal Age is also thought to have a significant influence on children's language outcomes, in that older mums tend to have completed more years of education, as well having better mental health (Zemencuk et al., 1995), greater emotional support (Moore & Brooks-Gunn, 2002) and knowledge about developmental milestones (Fergusson & Woodward, 1999). The benefits of being an older mother appears to taper off at around the age of 45, where children's language outcomes are similar to those with mothers under the age of 20 (Dunatchik et al., 2018), with some of the aforementioned explanatory factors, such as better mental health and emotional support, perhaps depleting once a certain age is reached. A similar U-shaped relationship is thought to exist between paternal age and children's language outcomes, with the youngest and oldest at the greatest risk of poorer outcomes (Malaspina et al., 2005). Parental age is also linked to income, in that older parents are more likely to be in a more financially secure position.

The influence of parental age and education on children's language outcomes, via knowledge gained, emotional support, financial security or otherwise, is necessarily qualified by other factors. For example, a parent's personality and parenting style will affect the way they interact with their child regardless of age or education. Similarly, the amount of time each parent spends with their child is also important, as presumably any potentially positive influence reaped from a highly educated parent will not be realised if the parent spends very little time with their child (although the influence of education could also be felt via other routes, such as childcare decisions). Again, it is also important to acknowledge that whilst higher levels of education are associated with the ability to provide an

optimal verbal environment for a child, for example by using a diverse range of vocabulary and reading a variety of books and nursery rhymes – it is clear that a high level of formal education is not a mandatory requirement for such behaviour.

Changing Demographic Trends: The Growing Role of the Father in Caregiving

Changing demographic trends regarding work and family life are an important consideration when considering the relative influence of mothers and fathers – through their age, education or otherwise. Historically, research on parental influence has tended to focus on mothers, who were traditionally more likely to play the role of carer. Yet change has been brewing over recent decades; Pleck (1997) noted that paternal involvement had increased over the previous three decades, both in absolute terms, and as a proportion of mothers' time. In the 21st century, fathers are now spending an increasing amount of time caring for their children (Burgess & Goldman, 2022), a shift that could result in a new wave of research focusing on the relevance of fathers to children's language outcomes. A report by the Fatherhood Institute, a charitable organisation interested in the role fathers play in their children's development, draws on 787 records including journal articles and reports, and incorporates data from three large population studies, including the Millennium Cohort Study (MCS) and the Avon Longitudinal Study of Parents and Children (ALSPAC). The report suggests that fathers' participation has increased dramatically over time: ALSPAC data in the early 1990s indicated just 39% of cohabiting fathers were regularly involved with feeding their infants, and 32% with bathing (with 28% playing no active role at all), yet by the early 2000s, MCS data indicated that 53% of fathers were now involved with feeding their infant at least once a day, and almost none reported no involvement at all (Calderwood et al., 2005, as cited in Burgess & Goldman, 2022).

But to what extent does a greater quantity or frequency of involvement from fathers' influence children's outcomes? Flouri & Malmberg (2012) found father involvement (measured via items such as frequency of feeding, bathing, playing) to be negatively associated with later emotional symptoms, although assessing the impact of fathers' involvement with their children is problematic, as estimates of quantity of involvement can be prone to error or bias (Pleck, 1997), and there are a number of confounding and interacting factors which may be at play. Research which only focuses on quantity or frequency of involvement also fails to consider elements of quality within that involvement – for example elements of responsiveness, stimulation and positive affect (as per Volling & Belsky's measure of father-infant interaction (1991), as cited in Pleck, 1997).

Particular features of fathers' involvement have been linked to language outcomes specifically. For example, in Tamis LeMonda et al. (2004), 111 low-income fathers were observed interacting with their children at 24 and 36 months old, with the aim of investigating the influence their interaction had on language outcomes, as well as comparing this to the mothers' influence. Both parents' sensitivity, positive regard and cognitive stimulation predicted later receptive and expressive language scores (suggesting similarities in parenting and influence on child outcomes), but only fathers' education and income predicted receptive and expressive language scores, as well as the former also predicting the quality of mother-child interactions. This latter point suggests that fathers had both a direct and indirect influence on child outcomes, not all of which was linked to their own physical involvement. Pancsofar and Vernon-Feagans (2010) also looked at low-income fathers and found that a father's education and his vocabulary use during shared storybook reading with his 6-month-old significantly predicted expressive language scores when the child was 36 months old, even when key

demographic variables, maternal education and maternal language input were controlled for. Conica et al. (2020) similarly found results which suggested a unique and important role for fathers, in that only a father's repetition (of their 2-year-old's utterances during a 10-minute triadic play interaction with the mother) was significantly related to the child's vocabulary diversity at age 4, possibly due to fathers' tendency to make clarification requests in question format.

In summary, the combination of changing demographic trends resulting in greater father involvement, and individual studies which have shown that particular elements of fathers' involvement are linked with child's language outcomes, suggests that the extent to which fathers caregiving plays a role in children's language development is a variable that is worthy of further investigation.

Number of Children in a Household and Birth Order

The Resource Dilution Model (Blake, 1981) posits that as family size increases, the resources available to each child deplete, including the time parents can spend interacting with each child, affecting their ability to construct language knowledge from verbal input. This assertion is generally supported by the wider literature, with siblings from larger families tending to have poorer language outcomes (Laing & Bergelson, 2024; Prime et al., 2014; Moss & Washbrook, 2016; Muñez et al., 2022). The assertion is also supported and somewhat entangled with research finding that firstborns (thus with no sibling competing for parental attention, at least for a while) are more proficient at acquiring language than subsequent siblings (Hoff-Ginsberg, 1998), although this finding was disputed by others (Pine, 1995; Rowe, 2012). It should also be borne in mind that the dilution of resources apparent in larger families as suggested by Blake (1981) has the potential to be mitigated by other factors. For example, where a parent has more children and therefore less time available for each one individually, this could be mitigated by having a good level of family and social support, with members who could step in to provide some of the care and interaction required, or indeed a larger family with a higher income also have the option of paying for additional childcare.

Gender

Gender is commonly included in large and small-scale studies investigating the influences on children's language outcomes, and as such has consistently demonstrated a robust relationship (Law et al., 2019; Dunatchik et al. 2018; Moss & Washbrook, 2016), with girls tending to have more favourable outcomes than boys. This gender gap is evident at ages 5, 7 and 11, with earlier outcomes likely contributing to later ones (Moss & Washbrook, 2016). Explanations for gender differences in language outcomes vary but include the idea that language learning tends to be promoted by other specific behaviours that girls tend to have in abundance – for example self-regulation, self-confidence (Adlof et al., 2010), task persistence and attentional control (Matthews et al., 2010). Another related explanation is that subtle gender-related differences in the way parents interact with their children may inadvertently encourage or discourage certain activities, some of which may promote language learning or related skills (Moss & Washbrook, 2016).

Dummy Usage

Whilst in comparison with more common demographic variables such as income, gender and education, dummy or 'pacifier' usage might be considered, in the strictest sense, more of a parenting choice than a demographic one. Nonetheless, information about an infant's dummy usage provides useful information about the characteristics of a particular sample, and thus can be considered demographic for these purposes. Dummy usage has been associated with a range of both positive and negative outcomes for children, including the aiding of breathing in pre-term infants and dental malformations (Boshart, 2001). The latter has been stated as one factor which may explain a suggested link between dummy usage and language development in children, in that the where the dummy sits in the mouth may cause malocclusions which affect the articulation of fricatives and alveolar phonemes (Boshart, 2001). Further reasons for the suggested link are an increased risk of otitis media, an infection of the middle ear, which would likely influence the sound input received by the infant during a critical period and potentially reduce the amount of babbling (Shriberg et al., 2000). Finally, it has also been suggested that the physical presence of the dummy results in less opportunity for babble and early word imitation (Burr et al, 2021; Shotts et al, 2008). Less chatter from the infant may elicit less response from the caregiver, thereby initiating a downward spiral of less quantity and quality of input and smaller vocabularies (Huttenlocher et al., 2010).

In a review of the evidence regarding the influence of both feeding and non-nutritive sucking behaviours (NNS) on speech sound development, Burr et al. (2021) found the evidence to be inconclusive, with the most consistent finding being the association between NNS duration and speech sound disorders – where both NNS and speech sound disorders were operationalised in various different ways, and NNS going beyond just dummy usage. In Barbosa et al. (2009) both finger sucking and extended use of pacifiers beyond three years in a sample of 128 Patagonian children aged 3-5 years was associated with a three-fold increased risk in developing a speech disorder, and in La Prairie (2010), whilst there were no significant differences between the articulation abilities of groups of 2- to 4-year-old pacifier and non-pacifier users, the non-pacifier group scored consistently higher on the articulation tests. Shotts et al. (2008) also found no significant difference when investigating the relationship between prolonged pacifier use and articulation, however they noted that future research should consider both the frequency of use (alongside overall duration), as well as the time of day the dummy was used. Both of these factors are likely to be important with respect to the aforementioned explanations for the link between dummy use and language difficulties (dental malformations, a higher risk of otitis media and less opportunity for babble and imitation), and whether or not such things come to fruition or not – for example a child who is reported to have used a dummy for over three years, but only for half an hour a day at bedtime, is less likely to be at risk of dental malformation than a more frequent user. Strutt et al. (2021) addressed some of these suggested gaps, measuring both duration and frequency of dummy usage of 100 twenty-four to sixty-one-month-olds, and found that most speech outcomes (including percentage of consonants, vowels and phonemes correct, as well as error type: age-appropriate, delayed or atypical), were not significantly associated with dummy usage, with the exception being the relationship between frequency of daytime use and atypical errors in younger children. This suggests that the link between dummy usage and subsequent vocabulary development may be worth investigating further, despite (or because of) the noted discrepancies.

Intersectionality of Risk Factors

Intersectionality refers to the interconnected nature of social categories and identities, such as race, class and gender, particularly within the context of inequality and disadvantage. Such a view proposes that these contexts are explained at the intersection and simultaneous occurrence of, for example poverty and lower levels of education, and thus cannot be fully explained by considering factors in isolation (Gross et al., 2016). This acknowledgement of intersectionality and simultaneous occurrence between risk-factors is demonstrated by the noted importance (above) of gender as a predictor of children's language outcomes: clearly not all boys are at risk of or end up with poorer language outcomes than all girls – gender must therefore be considered alongside other factors. Existing research tends to corroborate the intersectionality of gender and SES. For example, Barbu et al. (2015) tested 262 two to six-year-olds across four different groups (structured by gender and SES) on their acquisition of frequent phonological alternation in French, a phonological skill relating to the production of consonants between two words. Their results indicated that boys from a low-SES background had the lowest performance on this early language skill when compared to boys from a high-SES background, and girls from high and low-SES backgrounds (although the latter scored between the low-SES boys and high-SES children). The researchers suggested that although the interaction between child sex and SES was only marginally significant, it was still important for both factors to be considered together. Explanations provided for the differences between the groups included the idea that boys may be more influenced by SES than girls due to being more vulnerable to adverse circumstances, or, alternatively, that girls were more socially and emotionally responsive than boys, and thus benefitted more from the available language input.

The findings in Barbu et al. (2015) were somewhat consistent with those of Hammer et al. (2017). In this study, analysis of data from the Early Childhood Longitudinal Study (n=9,600) showed that the risk of being a late talker was significantly associated with both being a boy and having a low SES, although a number of other risk factors were also identified. Indeed, the intersectionality of gender and SES and the subsequent influence on language skills is not necessarily straight forward, as demonstrated by Lee and Al Otaiba (2015). Here, the researchers also examined how early literacy skills (in this case alphabet knowledge, phonological awareness and spelling) differed between groups (structured around gender and SES, as in Barbu et al.,2015), but a slightly more nuanced interaction between SES and gender was found. Boys and girls from low-SES backgrounds scored less on all three constructs than girls from high-SES backgrounds, but only on some constructs when compared to boys from high-SES backgrounds. Girls from low-SES households outperformed boys from low-SES households on two out of three of the measured constructs, indicating a relationship between the two variables, but one that isn't necessarily consistent across different aspects of language.

Another useful context within which to consider how different factors may intersect is the home learning environment (HLE), the quality of which has been consistently linked with children's language outcomes (Law et al., 2017a; Roulstone et al., 2012). The HLE consists of verbal input from caregivers, but also shared reading, age-appropriate toys, library visits, and attendance at childcare facilities (Roulstone et al., 2011). Poverty may inadvertently influence a parent's ability to provide an optimal verbal language environment for the reasons already discussed earlier in this chapter, but, simultaneously, there will also likely be less money for books, toys and expensive trips out. Lower income communities are also less likely to be able to access quality childcare provision (Pollard et al.,

2023), and libraries, whilst free at the point of use, may incur transport costs or other logistical difficulties. Poverty may also occur simultaneously with lower levels of parental education and larger households, both of which have the potential to influence the HLE, as discussed above when considering these variables separately. Thus, when considering poverty as a risk factor for language difficulties, it is important to acknowledge, in a holistic sense, the myriad things with which poverty may interact (not just the poverty in isolation), as well as noting that poverty is not in itself a guarantee of adverse outcomes, as gender is not.

Summary: Demographic Factors associated with Children's Language Outcomes

Numerous demographic factors have been linked with child language outcomes, and it is thus unfeasible to review each and every one here. Using Dunatchik et al.'s (2018) population-level study as a starting point, several commonly occurring factors have been explored above, including SES, parental age and education, fathers' role in caregiving, size of household, dummy usage and gender. This is not intended to suggest that these are the only factors involved, or that developmental and genetic mechanisms will not also be at play – simply that they may be important ones worthy of further investigation. Regardless of their probable importance, uncertainty surrounds the size and character of the role these factors play, and how they might be interrelated.

Ultimately, demographic risk factors are useful tools for identifying children who may be at risk of language delay, particularly when few alternative red flags exist. Yet when uncertainties persist about the significance of such factors, or when they are used as a blunt tool for identifying candidates for intervention, (for example using one factor in isolation to predict a child's language outcomes, or identify those at risk) further research is required to construct a more finely-grained approach to identify those who may need extra support.

The next chapter will explore the vocabulary gap, making a number of predictions about the characteristics of the words and word types which may reside within it, before going on to outline the rationale for this thesis.

Chapter 3: The Vocabulary Gap

Chapter 1 provided some general context regarding the importance of children's language skills, and specifically their range of vocabulary, before they start school. It noted the prevalence of language delay and language difficulties in general, as well as introducing the term 'vocabulary gap' and delving into some different theoretical perspectives regarding how children learn language, acknowledging the role of genetic and biological factors before focusing on the child's environment. Chapter 2 then went onto focus upon a child's verbal environment, particularly the quantity and quality of verbal input they receive and how this is related to their language outcomes. In the second part of the chapter, a number of demographic factors were reviewed due to existing evidence of their association with children's language outcomes, probably via their influence on a child's verbal environment.

The current chapter will return to the concept of the vocabulary gap, focusing on research which allows for tentative predictions to be made about the gap's qualitative content – in other words, what types of word may be missing for those children with smaller vocabularies. It will then provide a rationale for the current thesis.

Terminology

Usage of the term *gap* in promoting a deficit perspective of children's development has been acknowledged in Chapter 1, and yet there are further complexities with the term that must also be considered. *Gap* is often used by the media, politicians and academics when discussing children's capabilities, but it is important to note that there can be subtle differences in the type of gap being referred to. The '30-million-word gap' was a phrase coined by Hart and Risley (1995) and refers to a variation in parental verbal input (the words children hear). This variation in input then (according to interactionist theories of language acquisition) subsequently influences the number of words a child knows and uses themselves, resulting in a vocabulary gap, which refers to the differences between children's vocabularies (with some children knowing and using more words, and some less). The current thesis is concerned with the latter vocabulary gap – the differences between children's vocabularies, what form these differences may take, and ultimately, how to equalise these differences.

Cross-cultural, Historical and Broader Political Considerations

A wide variety in language proficiency amongst young children is not an issue confined to the UK, with evidence of similarly-defined vocabulary gaps in Canada, Australia and the US (Bradbury et al., 2015), Ethiopia, India, Peru and Vietnam (Boo, 2014), although the issue may not be as pronounced in countries where more uniform access to quality education and healthcare exists, such as Australia (Tucker-Drob & Bates, 2016). It is also informative to consider the historical nature of the vocabulary gap alongside achievement gaps in general. On the one hand, some evidence suggests that gaps may be declining, and equality improving (Ballarino et al., 2009; Breen et al., 2009). Many of the factors that have been associated with potentially adverse outcomes for children (e.g., poverty, poor mental health or parental education) are not newly arising societal issues, and the social, political and economic systems that have existed in previous years suggest much greater inequalities in the past than those that exist today. This is particularly true of wealth: in 1900-1910 the wealth of the top 10% of property owners was 94%, as compared to 48% for the same group in 1990 (Szreter, 2021, p.2).

Although this proportion has risen slowly since 1980, wealth inequality remains much lower than what it has been in previous years (Szreter, 2021). Indeed, research from the 1970 British Cohort Study found that children growing up in poverty in the 1970s were much more likely to score below their more advantaged counterparts, and also less likely to catch up (Feinstein, 2003).

On the other hand, survey evidence collating teachers' opinions suggests that the majority of primary and secondary school teachers believe that the vocabulary gap is widening (Harley et al., 2018), and there are numerous studies which testify to the ongoing stability of inequalities in educational attainment in general, particularly in industrialised nations since the early twentieth century (Shavit & Blossfeld, 1993). Indeed, if poverty is one of the demographic factors associated with language difficulties (even if this is a complex relationship, as discussed in Chapter 2), then any worsening in absolute poverty levels would suggest that such difficulties will indeed get progressively worse, unless sufficient countermeasures are introduced.

It is also worth considering that shifts in the UK education system have made any gaps appear more pronounced. Over the years, the UK education system has become more standardised, test-centred and accountable, beginning with the introduction of the National Curriculum in 1989, and the gradual introduction of SATs for KS1, KS2 and KS3, as well as the common use of similar but non-statutory tests at the end of each year (Poet et al., 2018). Whilst amendments have been made over the years, the direction appears to unambiguously in favour of testing, with the Year 1 phonics screening check introduced in 2012, and the reception baseline assessment in 2021, with varying levels of accountability for each test in terms of publication of results, and repercussions for Ofsted gradings. The benefits and drawbacks of this policy of regular statutory testing in schools is beyond the scope of this thesis, however of relevance here is how such a policy may have signposted language and achievement gaps more effectively, making them appear to get worse, when in actuality they were the same as they always were, or possibly even improving slightly.

In summary, opinion seems to be split on whether gaps and inequalities are improving or declining, and the absence of concrete data in the form of longitudinal datasets going back many years and using equivalent metrics for language and poverty make assertions either way difficult to make. Ultimately, the fact that the language gap exists is not negated by the direction of its movement, or that short-term trends may be susceptible to fluctuation regardless of long-term trends - but such movement may inform both the priorities of policymakers in the long term, and the effectiveness of interventions in the short term.

What does the Language Gap consist of?

Despite the discrepancies over what component of quality is most influential over vocabulary development, what exact mechanisms are responsible, and what child, family and parenting factors are most salient in predicting how a child's vocabulary develops, and whether or not they will experience delay or difficulties in general, the existence of a vocabulary gap, in terms of the quantity of words known or produced by a child, is very much supported (Huttenlocher et al., 2010; Fernald et al., 2013). Nevertheless, the nature of this gap has not been well studied, and at present it is not clear whether the variation is simply due to differential rates of learning along the same trajectory, or represents the difference between lexicons which are travelling along different trajectories, or are different in style (Frank et al., 2021). Regarding the latter, for example, it is conceivable that children

with persisting speech-motor difficulties will follow a different path of vocabulary learning to those who do not, in that they may have a continuing preference for words which are easier to pronounce. This suggestion is supported by Frank et al. (2021), who investigated other possible sources of individual variation in style of vocabulary learning, including referential versus expressive approaches to learning, and the effects of the so-called 'vocabulary spurt.' With respect to the former, Frank et al. (2021) confirmed that learners who were labelled 'referential' tended to have more nouns and larger vocabularies, whereas 'expressive' learners had fewer nouns but more syntactically complex speech, thus indicating a potential source of differences outside of rate alone. With respect to the vocabulary spurt however (referring to a proposed acceleration in word acquisition around the time of a child's first birthday), the researchers found no evidence within their CDI data, and as such suggested it could not be a source of variation having the potential to vary in terms of the onset, type and duration of accelerated growth.

Whether it be due to rate or style, information pertaining to the vocabulary gap that goes beyond the mere fact that it exists is important, because a large amount of effort and money is currently directed at 'closing the gap', particularly the language gap, which, as noted in Chapter 1, is linked to numerous academic and well-being indicators in later life, including academic results (Snowling et al., 2011), job prospects, (Gascoigne & Gross, 2017), social and emotional well-being (for example an increased likelihood of psychiatric disorders; Beitchman et al., 2001), youth offending (Royal College of Speech & Language Therapists, 2012), and even life expectancy (Gilbert et al., 2018). If the knowledge concerning what the vocabulary gap consists of does not yet exist, or is under-defined, then this raises the question as to whether interventions designed to help children improve their vocabulary (and therefore narrow the gap) are focusing on the right thing.

The lack of research regarding the nature of the vocabulary gap, and the sorts of differences which exist between children with large and small vocabularies, make it difficult to conjecture what these differences might be. One way of circumventing this difficulty, based on the presumption that children with smaller vocabularies are simply further behind on the same developmental trajectory than those with larger vocabularies (as opposed to a trajectory which is different in style altogether, as discussed above), is to derive predictions about the nature of the gap using research on the age of acquisition (AofA) of words. If such research were to show that words of a particular type or with particular characteristics (for example long words, or infrequent words) tend to be learned later, then this would suggest that such words may form part of the vocabulary gap, and be missing from the lexicons of those with smaller vocabularies. Another more complex strategy for making predictions about the nature of the vocabulary gap would be to look at differences in child-directed speech (CDS) amongst particular demographic groups which have been linked to different vocabulary outcomes for children, and corresponding differences in child vocabulary. This strategy is briefly considered later in the chapter, however due to the current lack of research focusing on the qualitative differences in child language outcomes in terms of the composition of their lexicons, it was determined to have limited utility in predicting what words may form part of the language gap for the current research.

Using Age of Acquisition Research to make Predictions about the Vocabulary Gap

Age of acquisition refers to the age at which a word is typically learned, and thus can provide useful information about the order in which children tend to acquire different word types. Assuming a similar trajectory of vocabulary acquisition for all children, it therefore appears plausible that words

which are acquired later are more likely to reside within the vocabulary gap and be missing from the lexicons of those children with smaller vocabularies. AofA data can work alongside theories of word learning, including the ECM, in that explanations for why particular word types tend to be learned earlier or later can be sought via such theories, for example word types which tend to be less concrete in nature may be learned at a later stage due to requiring the use of more sophisticated cues which have not yet developed.

The MacArthur Bates CDI data repository provides access to norming data for the typical age of acquisition of different words: 680 items for children aged 1:4 to 2:6 (Dale & Fenson, 1996). This age of acquisition is calculated to be the first month in which 50% of children within the norming study were reported to produce (or comprehend) the particular word (Goodman et al., 2008). The CDI, as discussed in Chapter 1, is a parent report measure designed to measure both receptive and productive vocabulary from 8 months old in a convenient and inexpensive way across large samples of parents and different languages. Parent report is considered to be an accurate measure of children's knowledge, in that parents generally have considerable experience of their child's verbal behaviour across various different contexts and over time, behaviour which is less likely to be affected by the specifics of a single (or even multiple) experimental or observational contexts, for example a limited toy selection within a clinic setting, or shyness when being spoken to or observed by an unfamiliar researcher (Frank et al., 2021).

Despite these stated benefits, it is important to acknowledge the potential for parental bias in CDI reporting, in that a parent's judgment of whether their child knows a particular word is likely to be based on a complex set of factors including both specific recall of usage and their own subjective opinions or how 'difficult' the word is (Frank et al., 2021). These biases could involve both under-estimations and over-estimations of a child's ability, so it is difficult to ascertain the potential effect they may have on derived AofA data, and reporting difficulties appear to be more apparent with comprehension rather than production, the latter being of interest here. A further potential limitation of the CDI (and therefore data derived from it) is the fact that it functions as an index of children's vocabulary, rather than a full and complete measure (e.g., 680 words for the CDI-Words & Sentences). If this index was not an accurate reflection of the types of words children know, and an accurate reflection of the composition of children's vocabulary, then whilst the AofA data for these words would still be broadly correct, the data would not necessarily be focused upon the words which children knew and were using, thus any predictions based upon them would be flawed. Nonetheless, there exists no complete measure of a child's entire lexicon or perfect way to collect children's vocabulary data, with other types of research also subject to limitations (see Chapter 6 for further discussion on the limitations of naturalistic language samples). Research which uses CDI data, (and particularly the AofA norming data which is derived from it; Dale & Fenson, 1996), is thus deemed to be a reliable evidence base for making predictions about children's vocabulary acquisition, either on its own or alongside research evidence which uses alternative approaches for studying how children learn vocabulary. Linguistic variables which are commonly found to influence AofA (such as word length and lexical category) are therefore reviewed below, with a view to making predictions about the order in which different word types are acquired, and therefore the types which may form the body of the word gap.

Word Frequency and Rare Words

One common theme to emerge from the literature is word frequency; the more often a word is heard by a child, the more likely it will become known and used by them (Swingley & Humphrey, 2018; Huttenlocher et al., 1991), with words heard less commonly being less likely to be acquired early on. Very uncommon words (often termed *rare* or *sophisticated*), are likely, by definition, to be infrequent in nature, however the differing ways which the measures can be operationalised in studies suggest that 'rare' and 'frequency' may capture slightly different aspects of a word's commonality. Rare words represent only 2-6% of maternal talk, and yet exposure to such words has been shown to be particularly predictive of later vocabulary size in general (Rowe, 2012; Weizman & Snow, 2001), as measured by standardised tests (for example in Rowe, 2012, caregiver use of rare words with their 30-month-olds significantly predicted receptive vocabulary scores on the PPVT at 42 months, with r^2 = .58). These studies show that increasing the frequency of otherwise infrequent words is beneficial for later vocabulary, but they do not pinpoint exact vocabulary outcomes, only an influence on general vocabulary size – indeed those children with large existing vocabularies may find it easier to pick up all types of words, including rare words. Hence, the assumption of frequency affecting language outcomes may be somewhat over-simplified; as we have seen, the influence of frequency may be qualified by other factors, such as whether or not the words are spoken frequently during periods of joint attention (Tomasello & Haberl, 2003) or contingent talk (McGillion et al., 2017), or embedded in useful and instructive interactions (Weisman & Snow, 2001), or if the word occurs in an utterance's final position (Naigles & Hoff-Ginsberg, 1998). The issue is further complicated by the fact that measures of word frequency often differ between studies – for example some use a measure of word frequency based on maternal CDS (e.g., from CHILDES transcripts), and others use more adult measures of word frequency (e.g., the British National Corpus).

As alluded to above, some studies have suggested a more complicated role for frequency in word learning. Goodman et al. (2008) argued that earlier studies were flawed in that they could not explain why specific highly frequent words were learned earlier than less frequent ones, as they tended to only correlate the amount of the mother's speech with the child's overall vocabulary size, or occasionally particular syntactic categories in particular linguistic communities. In addition, experimental studies, (e.g., Schwartz & Terrell, 1983; Rice et al., 1994), have tended to examine only one or two novel words, or in the case of Blackwell (2005) - a corpus analysis - single lexical categories. Goodman et al. (2008) built on these findings by examining a large number of words from the full range of lexical categories seen in early language in both production and comprehension, using data from the Wordbank Project (containing norming data on the age of acquisition for 562 words) and the CHILDES database (estimating frequency on parental usage of particular words from corpora), and found that higher frequency in parental input was only significantly associated with earlier acquisition within specific lexical categories – not with the lexicon as a whole (r=-.07; p= .055) and its influence also varied depending on developmental stage. For example, whilst there was a significant correlation between frequency and nouns throughout the development period studied, this correlation became stronger later on, after the first 100 words had been learned (r=.43, compared to r=.19 for the first 100 words), implicating the need for other explanatory factors regarding the noun bias present in the early stages of the majority of languages.

Other studies have also found qualifications to the frequency-ease of acquisition relationship. Stokes (2010) suggested that limited conclusions could be drawn about frequency as a predictor due to the high correlation found between frequency and phonological neighbourhood density – in this study, word frequency was not significantly correlated with vocabulary size for those children with comparatively smaller vocabularies. Brent and Siskind (2001) created and studied a corpus of eight mother-child dyads and found total frequency was not a significant predictor of how early a word would be acquired, regardless of which way the child's vocabulary was measured, via the CDI or analysis of children's productions (r=0; p=.856 and r=0; p=.181), but how often a word was said in isolation was (r=.07; p=.002 and r=.07; p=.016), demonstrating the importance of segmentation ease – something that becomes easier as the lexicon develops (Alvater-Mackenson & Mani, 2013). Swingley and Humphrey (2018) used the same data to re-examine that conclusion, but this time included a wider range of predictors to decipher whether frequency of isolation was hiding behind some other factor, such as length, concreteness, phonological simplicity or a tendency for hyper-articulation. On the contrary, in this case the researchers did find overall word frequency to be the most consistent and important predictor of productive (and receptive) vocabulary (r=.28; p<.001 and r=.48; p<.001), possibly because log frequency rather than raw frequency was calculated here (in contrast to the Brent and Siskind study). How often a word was heard in isolation also significantly predicted both productive and receptive vocabulary.

In summary then, word frequency and rareness – which are likely to be similar measures (but this depends on how each is operationalised within the study) – appear to play an important role in a child's vocabulary development, both in terms of how early that specific word or word type is acquired, and also via a general effect on vocabulary size. Support for this assertion is not universal however, and the relationship between frequency and children's vocabulary may also be qualified by other variables, such as lexical category or phonological neighbourhood density. Whilst different measures of word frequency are generally correlated (Stokes, 2010), variation in the way individual studies calculate word frequency (for example using estimates from CDS versus more adult-orientated estimates like the BNC) also makes ascertaining the overall influence of frequency on children's vocabulary via comparisons between studies more difficult, in that it is likely to have a moderating effect on the extent to which the influence of frequency can be observed. Furthermore, if the measure of frequency is not based on the CDS a child hears from their own parent, then regardless of which measure this is, there will not be a complete match between words which are frequently heard, and those which are considered frequent by some external measure. This mismatch may mean that 'frequent' words are not necessarily the driving force behind children acquiring particular word types. Nonetheless, it is clear that word frequency, in some guise, has a role to play in children's vocabulary acquisition, and thus is a worthy predictor to include here. It is likely that less frequent or rare words will reside within the vocabulary gap, as children with smaller vocabularies may have experienced less exposure to these words.

Lexical Categories

The vast majority of languages appear to exhibit some degree of noun bias, referring to an over-representation of nouns in children's early vocabularies when compared to other lexical categories (Frank et al., 2021). With a sample of 1130 children aged between 1:4 and 2:6, Bates et al. (1994) analysed CDI data and found the proportion of nouns in children's productive vocabularies to

be the largest when compared to both predicates and closed-class words, with a mean of 41.9% across the age range, although this proportion varied depending on the size of a child's vocabulary, with the proportion of nouns peaking and levelling off around the 200 word mark. Also using CDI data, Frank et al., (2021) confirmed the existence of a noun bias in almost every language examined (including different versions of English), with the effect more pronounced for production than for comprehension, and the bias against the early acquisition of function words being the largest and most consistent when the entire sample of languages was considered. These findings suggest that nouns may be somehow easier to acquire than other lexical categories, including predicates and function words, possibly because nouns are more likely to be refer to common everyday objects that children are interested in, or due to their morphological simplicity when compared to other word types, including verbs (Frank et al., 2021). Syntactic category may also be influential when interacting with other variables, as was the case in Goodman et al. (2008), where frequency was found to be important only when considered within each individual syntactic category, rather than as an overall effect.

Although the study was concerned with comprehension as opposed to productive data from the CDI, the importance of lexical categories was nonetheless confirmed by Braginsky et al. (2016). Here, the authors analysed a subset of data from The Wordbank Project (containing CDI data for 40,000 children across 14 countries), thereby estimating the age of acquisition of 300-400 words for 9200 children across seven languages, and calculating corresponding frequency estimates of these words using the CHILDES database. Results confirmed and expanded on previous findings, in that the study was able to focus on numerous different predictors (including frequency, mean length of utterance (MLU), length, concreteness, valence, arousal and relevance) simultaneously so that their relative influence could be studied. Importantly, the predictors were found to vary in strength across different lexical categories, with frequent concrete nouns learned earlier, yet for predicates concreteness was not as important, and for function words the length of utterance was most predictive in how early such words would be acquired. The authors suggested that nouns may be acquired first because they usually have easier referents in the form of everyday objects, which tend to be more concrete or 'imageable' (Hansen, 2017), although this was not always the case, particularly concerning abstract or proper nouns. Verbs, on the other hand, have a more complex function which often denotes the relationship between objects (e.g., *the girl kicked the ball*). Even if the verb resides within an intransitive form without a direct object (e.g., *the soldiers arrived*), such a form still requires a multi-word utterance to communicate meaning effectively (as opposed to nouns, which can make sense when uttered in isolation). Children tend to start using multi-word utterances later on, once they have an established bank of vocabulary, at around 18-24 months, and so early speech may be more suited to words (i.e., nouns) which can more easily communicate meaning when said as a single word utterance.

In summary then, nouns, particularly concrete nouns with easily identifiable referents (Baldwin, 1995), appear to be more easily acquired than other word categories, although this is not necessarily because they are more frequently heard (Goodman et al., 2008). Should this be the case, other lexical categories, or subgroups of these, may be more likely to be find themselves located within the language gap. Lexical categories may also be important as moderators of other effects (Braginsky et al., 2016). For these reasons, lexical category is an important predictor of vocabulary acquisition, and will be included here. It follows that if nouns are easier to acquire than other word

types, it will be these other word types that are more likely to reside within the vocabulary gap, and be missing from smaller vocabularies.

Concreteness

A further common theme arising from the literature is concreteness, referring to the word's tangibility versus abstractness (Jones & Brandt, 2019). Words have been shown to be easier to acquire if they are more concrete, although this may be qualified by lexical category (Braginsky et al., 2016; Goodman et al., 2008), developmental stage (Jones & Brandt, 2019) or whether the outcome measure is receptive or productive vocabulary (Swingley & Humphrey, 2018). In Jones and Brandt (2019), a positive relationship was found between a word's concreteness and the age at which it was produced and comprehended (Figure 1, p.1266), with the association between concreteness and production increasing over time between 12–25 months (but not for comprehension). Conversely, in Swingley & Humphrey (2018), a word's concreteness was found to have a significant and positive effect on comprehension (r=.35; p=.01) but not production, possibly because concreteness is important for deciphering meaning, and referential learning, but once a word is understood, barriers to speaking a word may be more phonological, although of course the two will be linked. Bias in parental reporting was another reason put forward by the authors, in that parents are more likely to believe that their children know words relating to concrete rather than abstract concepts. In summary therefore, despite some discrepancies within the existing evidence, it appears likely that concreteness plays some sort of role in ease of word acquisition, even if this may be stronger for comprehension than for production, and thus is an important predictor to include here. If concrete words are in general easier to acquire, it thus follows that the vocabulary gap will have a higher proportion of abstract words.

Word Length and Pronunciation Difficulty

Word length and pronunciation difficulty are separate yet related variables that may affect how easy a word is to acquire, the former because longer words have longer phoneme sequences and therefore require a greater number of sub-lexical phonemes to be stored and then successfully retrieved from memory (Jones et al., 2021), and the latter because words that are difficult to pronounce may present articulation difficulties: certainly words that are acquired later have been shown to be more difficult to articulate as measured by speed of repeated articulation (Gerhand & Barry, 1998), although current evidence is lacking and ambiguous. Due to measures of word length and pronunciation difficulty tending to be operationalised by the same metrics, including number of syllables, characters or phonemes, all of which have been found to be highly correlated (for example Lewis & Frank, 2016, found the correlation between phonemes and characters to be r=.92, phonemes and syllables r=.89 and syllables and characters r=.87) – reviewing the evidence on how important these variables are is not straight forward.

A number of studies have found length to be an important factor for children's word learning, in that shorter words tend to be easier and therefore learned first. Frank et al. (2021) found that the number of phonemes (here used as a marker for pronunciation) was significant in that shorter words tended to be learned first, but only for productive vocabulary. The fact that this was not the case for receptive vocabulary might be explained by the fact that shorter words tend to be easier to pronounce, but this was not so important for comprehension. Roy et al. (2015) also found number of phonemes, here representing ease of production, to be predictive of word learning, although this was secondary

to other findings, and involved a single case study. Jones et al. (2021) measured length in terms of the number of syllables and found it to be an important factor concerning ease of acquisition, as the phonological knowledge required for shorter words with shorter chunks or phoneme sequences would necessarily be less, and more likely to have been encountered before, and therefore more likely to exist in a child's repertoire of existing knowledge.

On the other hand, Swingley and Humphrey (2018) indicated that number of syllables might not be important for word learning. Here, the researchers included a number of predictors of word acquisition in their initial model, including word frequency, frequency of word spoken in isolation, final or medial positions within the utterance, concreteness, word category, MLU, word duration and phonological variables such as PP and number of syllables - however a number of variables, including number of syllables, did not improve model fit and were not retained in the final model – indicating that number of syllables 'did not have significant predictive power' in this particular study (Swingley & Humphrey, 2018, p.1262). Braginsky et al. (2016) – using the same CDI dataset as Frank et al. (2021) but focusing on language comprehension, also found length (here measured in characters) to be a much weaker predictor of acquisition across seven different languages when compared to frequency, babiness (referring to how relevant a word was to babies), concreteness and MLU, and thus it was omitted from the final analysis.

In summary, despite the difficulties in reconciling the studies reviewed above, length and pronunciation have been shown to be important influencers of how and when children learn words, and thus will be included here. It is likely that words that are more difficult to pronounce will form part of the language gap, not only because they put pressure on a child's speech motor ability (Frank et al., 2021), but also because they are more likely to be absent from their parents' vocabulary, and thus the vocabulary they are exposed to. Longer words are also likely to form part of the language gap, as longer phoneme sequences require a larger store of sub-lexical phonemes to decipher the word (Jones et al., 2021). Having said that, the fact that that longer word length (be that measured in characters, syllables or phonemes) does not necessarily equate to greater difficulty in pronunciation (e.g., *sixth* is more difficult to pronounce than *difficult*, despite being shorter) is perhaps an indication that the two factors should be treated as separate variables and measured via different metrics in future, for example pronunciation difficulty could be measured via the presence of specific complex consonant clusters in words known to have a later age of acquisition.

Phonological Neighbourhood Density and Phonotactic Probability

Phonological neighbourhood density (PND), a measure of word similarity, refers to the number of words that differ to the target word by a single phoneme, hence a dense neighbourhood would imply a word has many neighbours, and is therefore phonologically similar to many words. Words with dense neighbourhoods have been established as easier to learn than those with a sparse PND (Storkel, 2004). Gathercole et al. (1999) suggested that this was due to the fact that cognitive demand was low when processing new words comprising of commonly occurring sounds, allowing detailed long term phonological word memories to be formed, which supported word production. Jones et al. (2021) suggested that words with high neighbourhood densities shared a greater number of sound sequences with other words than words with low neighbourhood densities, and therefore phonological knowledge is learned more quickly for the former than the latter.

A number of studies have supported the assertion that words with dense neighbourhoods tend to be easier to learn, and thus learned earlier. Jones and Brandt (2019) used a subset of CDI data (300 British-speaking children aged between 12-25 months) alongside multiple other data sources, including frequency estimates using CDS from the Manchester Corpus in CHILDES, to estimate the strength of PND relative to other predictors, including frequency, length, valence, concreteness, babiness, arousal and dominance. Here, the same result was found (in that words with a high neighbourhood density were learned more easily by children), but only for production and not comprehension, and with stronger effects in younger children. The authors suggested that words with a dense neighbourhood were easier to memorise in a phonological manner, and could therefore be more easily produced. Such phonological memorisation was not required for the comprehension of words however, as children could still recognise and respond to words even when they had imperfect phonological representations of them. A relationship was also found between word length and PND, in that shorter words tended to have a denser neighbourhood, and this could be the underlying reason for their ease of acquisition.

The notion of words with denser neighbourhoods being easier to learn was supported by Stokes (2010), although its influence may vary depending on a child's stage of development. Here, CDI data for 222 British-speaking children showed that PND was more important for those children with smaller vocabularies, in that they appeared to have more words with higher neighbourhood densities, but lower frequencies, whereas for those children with large vocabularies, the opposite was true. The authors suggested that words with higher densities and lower frequencies represented lexical difficulty, and such words tended to be spoken with expanded vowel spaces and increased duration, hence maximising listener perception.

Conversely, not all researchers have found PND to have a significant effect on word learning (Swingley & Humphrey, 2018), or not in the expected direction anyway (Swingley & Aslin, 2007). In the latter study, the researchers conducted experiments in which seventy 18-month-olds were exposed to pairs of novel words consisting of either familiar sound patterns (a dense neighbourhood) or non-familiar sound patterns (a sparse neighbourhood) - with children performing better on the words that were non-neighbours, with the difference between the two conditions (Experiment 2) being statistically significant ($t(26)=1.98$, $p<.05$). The authors suggested that close neighbours can create a hindrance to word learning where the words are too similar, with the known familiar word interfering with acquiring the new but similar word.

A factor related to PND and also thought to influence word acquisition is Phonotactic Probability (PP). PP refers to the likelihood of a particular sound sequence within a word occurring again, and its effects can sometimes be difficult to disentangle from those of PND (Swingley & Aslin, 2007), with PP thought to influence activation of phonological representations, and PND thought to influence activation of lexical representations (Storkel et al., 2010). In Storkel et al. (2010), the researchers aimed to differentiate the effects of PP and PND on naturalistic word learning on two separate groups of pre-school children – 34 of whom were typically developing, and 20 of whom had some phonological delays. Whilst PP and PND have been shown to be correlated variables, in that common sound sequences tend to reside in dense neighbourhoods (Storkel, 2004), Storkel et al. (2010) were able to show, in fact, that words with either a combination of common PP and dense PND (more 'typical' words), or rare PP and sparse PND (more 'distinctive' words), were the easiest to acquire for

typically developing children, with only the latter combination significant for delayed children. The authors explained this apparent paradox by suggesting that different mechanisms were at play at either end of the scale, with distinctive words being capable of triggering word learning more efficiently, and typical words being easier to retain in working memory, hence making both typical and distinctive words the easiest to learn.

Using a computation model able to learn chunk-based phonological knowledge in conjunction with corpus data from the Manchester Corpus available via the CHILDES database, Jones et al. (2021) found that length was the deciding factor when it came to PP; words with a common PP were learned more quickly, but only once word length was controlled. When length was not controlled, it was words with a rarer PP that tended to be learned more easily, as they tended to be shorter. Shorter words were more likely to be learned before longer words, as the phonological knowledge required for shorter words with shorter chunks or phoneme sequences will necessarily be less, and for high frequency words the process will be even quicker. The authors suggested that this effect of length on PP provided an explanation for inconsistencies in the literature, in that studies that found a high PP advantage were usually focusing on words of a particular length, and those that found the opposite were usually testing words of all lengths. Bearing this in mind, they questioned the worth of using PP and PND as markers of phonological knowledge and word learning in general at all, with the ultimate marker of how easily a child learns words being the presence of prior and accumulating phonological knowledge, of varying 'grain size' (Jones et al., 2021, p.11).

In summary, the majority of the evidence points to the conclusion that words with a dense PND are easier to acquire (in that they have more phonologically similar word neighbours), but this may depend on the existing lexicon of the child (Stokes, 2010), and has not been found universally (Swingley & Aslin, 2007). Inconsistencies are also apparent in the evidence surrounding PP (the likelihood of a particular sound sequence occurring again), as whilst PP and PND are related variables, with common sound sequences tending to reside within dense neighbourhoods (suggesting common sound sequence will be easier to learn), research suggests that when a word with a low PP value is combined with a sparse PND (thus making it more distinctive), this is more predictive for children with smaller vocabularies (Storkel et al., 2010). Both PND and PP may be confounded by word length (Jones et al., 2021; Jones & Brandt, 2019), with some commentators suggesting they may have limited worth as predictors of word acquisition. Nonetheless, the level of ambiguity around their importance is justification of further investigation, and thus their inclusion as predictors of vocabulary in the current study.

Other Possible Predictors of Age of Acquisition and the Vocabulary Gap

In addition to the common themes discussed above, numerous other suggestions have been made regarding what makes a word easier to acquire, and therefore, by implication, less likely to be present in the language gap. These suggestions include: babiness (Tardif et al., 2008; Perry et al., 2015), words that tend to be hyper-articulated (Swingley & Humphrey, 2018), words that tend to appear in shorter utterances or at the end of utterances (Jones & Brandt, 2019), words with wider semantic networks in the immediate learning environment (Hills et al., 2009) and words that appear in more distinctive spatial, temporal or linguistic contexts (Roy et al., 2015). In the latter study, Roy et al. (2015) employed an in-depth case study to show that these three dimensions of context – including the location in physical space, the time of day, and other words that appear nearby in the conversation -

were stronger predictors of AofA than frequency, MLU or word length, and played a consistent role across syntactic categories. Despite their suggested importance, the evidence for these aforementioned suggestions is more limited and/or focused on analysis that goes beyond the word level (e.g., the word's position in an utterance), and thus they have warranted only a brief mention here.

The Potential for making Predictions based on Differences in Child-Directed Speech (CDS)

As noted earlier in this chapter, if, as is probable, the language gap represents a delayed rate of learning along a similar trajectory, where children with smaller vocabularies are simply further behind on the same path of acquisition, then AofA research can provide useful predictions about the order in which different word types or words with particular characteristics are learned. Despite this being the approach decided upon for the current research project, it is important to acknowledge that an alternative strategy could have been utilised to predict which words the vocabulary gap might contain in the form of predictions based on child-directed speech (CDS). Along these lines, differences in CDS amongst particular demographic groups which have been linked to different vocabulary outcomes for children could have been analysed and used to make predictions about any corresponding differences in children's vocabulary. This approach may have been particularly suitable if the language trajectories of those on either side of the gap were thought to be different in style, as discussed earlier in this chapter (see Frank et al, 2021), perhaps due to differences in speech-motor ability or caregiver verbal input (e.g., more directives over questioning and less contingent talk, see Leigh et al., 2011).

Despite the potential suitability of CDS data for making predictions about the differences in children's language and the vocabulary gap, studies which link particular linguistic characteristics of CDS to the same features in children's subsequent speech do not provide enough detail to make a full range of predictions (e.g., Naigles & Hoff-Ginsberg, 1998, found that relative verb frequency and diversity of syntactic frames used by caregivers were strong indicators of the child's early verb use; Hoff-Ginsberg, 1986 found that the average number of noun phrases per utterance was a positive predictor of the same measure in children's speech). This lack of detail, and the tendency for studies which investigate CDS to focus on the differing use of single linguistic features, rather than the overall characteristics and composition of CDS of different groups of parents that have been linked to different child vocabulary outcomes, excluded this option for the current research.

A further and perhaps more operationally significant barrier to using CDS research as a strategy for making predictions about the vocabulary gap is that there are many different demographic variables that have been associated with children's language outcomes, and upon which one should parents be grouped and differences in CDS analysed? Existing research looking at differences in CDS tends to focus on the previously established but somewhat blunt link between SES (particularly maternal education) and CDS characteristics (Hart & Risley, 1995). For example Hoff (2003) looked at properties of maternal speech including word tokens, types and MLU, and found them all to be correlated with maternal SES and subsequent corresponding child vocabulary. As noted in Dunatchik et al. (2018), there are numerous socio-demographic factors beyond the more obvious components of SES (such as education, income and occupation) that are associated with vocabulary outcomes for children, including parental stress, the HLE (home learning environment), parental aspirations and many more. With these myriad factors involved, it would present quite a challenge to

create groups of parents which met all these risk-factors simultaneously, yet to only analyse CDS differences based on one factor (e.g., SES, as above in Hoff, 2003) also seems unsatisfactory. Furthermore, there is also likely to be substantial individual differences within groups (however categorised) which could render any patterns meaningless.

In summary, CDS has been acknowledged as an alternative strategy for making predictions about the vocabulary gap, but for numerous reasons this strategy has not been employed for the current research. Firstly, existing research tends to focus on single linguistic features which do not provide enough detail to make a full range of predictions about the nature of vocabulary gap, and secondly, that there would be inherent difficulties in deciding how to categorise and match different parent groups (in order to analyse the characteristics and composition of their CDS) based on the many different risk-factors known to be associated with children's language outcomes. Nonetheless, the fact that some research shows qualitative elements of parental speech to be mirrored in subsequent child speech (Pine, 1992; Naigles & Hoff-Ginsberg, 1998; Hoff-Ginsberg, 1986; Huttenlocher, 2002) suggests that there is the potential to make predictions about the characteristics of children's speech, and the differences between their vocabularies, in future research.

Summary of Chapter 3

The vocabulary or language gap is the difference between the vocabulary size of those children with the largest and smallest lexicons, and is related to but conceptually different to the term *word gap*, which refers to differences in parental verbal input. The vocabulary gap is evident worldwide, and whilst the trajectory over time of achievement inequalities in general is debatable, 'gaps' are perhaps more clearly signposted by today's culture of statutory testing within schools.

The qualitative nature of the vocabulary gap, in terms of what words or word types tend to be known by those with the largest vocabularies but not by those with the smallest, has not yet been fully researched. Assuming that children with smaller vocabularies are simply further behind on the same developmental trajectory than those with larger vocabularies, it is possible to derive predictions about the nature of the gap using research on the age of acquisition (AofA) of words. Such research, as reviewed throughout this chapter, tentatively predicts that the vocabulary gap will consist of rare words, less frequent words, more abstract words, longer and more difficult to pronounce words, words which have a denser PND (more phonologically similar word neighbours) and a lower PP (less common sound sequences), and words which belong to a word class other than nouns. Complex interactions between these variables are likely, with other factors potentially at play, including a word's babiness or whether it tends to appear in more distinctive spatial, temporal or linguistic contexts (but limited evidence precluded these factors as predictors for the current study). Whilst it has been acknowledged that predictions, from an interactionist perspective, could also have been gleaned from research looking at qualitative differences in CDS, this strategy was discounted for the current research.

Thesis Justification

The language skills children develop in their first few years of life have important individual and societal consequences (Gascoigne & Gross, 2017; Field, 2010), and incorporate a number of different but related skills, including the diversity of words known and used by a child – their

vocabulary. The gap that exists between those with the largest and smallest vocabularies has been well documented (e.g. Huttenlocher et al., 2010), with the gap predicted to become more pronounced in the post-pandemic years (Tracey et al., 2022). Such a gap does not correspond well with society's professed aim of providing equal opportunity for all, and thus considerable funding has been (and will continue to be) made available to close that gap. Whilst adequate funding is a good start, the challenges involved in improving children's language outcomes appear to transcend financial provision.

Many intervention initiatives currently in circulation aim to 'close the gap', and yet whilst there is robust evidence that the gap exists (Fernald et al., 2013), the qualitative nature of the vocabulary gap, in terms of the word types or characteristics which may form the gap, remains unclear. This suggests that there is some uncertainty around the efficacy of such initiatives: if knowledge regarding which words or word types form the body of the gap is not yet known, then it must be conceded that the content of specific interventions designed to close the gap may be mis- or under-informed.

One way to investigate the vocabulary gap is utilise a corpus of children's talk, and to see which particular linguistic characteristics, as predicted by the AofA research reviewed in this chapter, are associated with vocabulary size, thereby revealing the differences between children's vocabularies, and particularly those with the largest and smallest vocabularies. Research Objective 1 will therefore be to explore the linguistic characteristics of the vocabulary gap, with a view to formulating a list of particular word types which are characteristic of the vocabulary gap. These word types will likely be missing from the lexicons of children with the smallest vocabularies for their age, and should necessarily form part of the content of interventions designed to close the gap.

Ascertaining the qualitative nature of the vocabulary gap is not the only challenge apparent here. As noted in Chapter 2, identifying children who are at risk of a language delay before they reach school is complicated, particularly if there are no language or behaviour-related red flags. In this way, demographic risk factors can be useful tools to identify children who might need extra support. Many such demographic factors have been associated with children's language outcomes (Dunatchik et al., 2018; Law et al., 2019), but there is uncertainty around their relative importance and the size and character of the role they play. The widely used current strategy of targeting children via their SES/income strata alone is over-simplified. Whilst it is unfeasible to review the relative importance of every demographic risk factor thought to be influential, further analysis of two existing datasets, using a variety of language measures and focusing on the reviewed factors in Chapter 2, is necessary to provide a more finely grained picture of risk. Research Objective 2 is thus to investigate the most salient demographic risk factors associated with smaller vocabularies at pre-school age.

The outcomes materialising from addressing Research Objectives 1 and 2 will then be utilised in the second phase of this project, which will be focused upon developing an intervention aiming to improve children's vocabulary, and with the potential to 'close the gap'. The research question for this final phase will be: how can a vocabulary intervention be effectively incorporated into an audio format? This phase will use Action Research methodology, and the context and rationale for its chosen approach will be considered in more detail in Chapters 7-9.

Chapter 4: Method (Phase One)

Research Objectives

The purpose of Phase One of this thesis was to investigate the expressive vocabulary differences evident in pre-school children in a natural setting, with a view to ascertaining whether vocabulary (measured in terms of lexical diversity) could be predicted by particular linguistic and demographic measures. Results obtained would then be used to form the basis of an intervention in Phase Two. Further details on the context of the research problem can be found in Chapter 1. This chapter will describe the research design in detail, as well as providing a justification for why such a design was appropriate to address the research questions at hand.

Research Paradigm: Pragmatism

The decision a researcher makes about which research paradigm to employ can have far-reaching implications for their work, from the phrasing of research questions to decisions about how the data should be analysed, and thus it is an important choice to reflect upon before further details regarding methodology are considered. A research paradigm is a set of common beliefs or world view about how problems should be understood and addressed (Kuhn, 1962), encompassing specific ontological, epistemological, methodological and axiological positions (Sandelowski, 2000). Pragmatism arose as a philosophy centred on the idea that human thoughts and experience are intrinsically linked to their actions, and therefore action is pivotal (Goldkuhl, 2012; Morgan 2014). As a research paradigm it is less concerned with the nature of reality and truth addressed by ontology and epistemology, but proposes that the researcher focus on the problem to be solved and then use the best range of methodological approaches available to investigate that problem from various perspectives (Dewey, 1910). Such an approach lends itself to solving practical problems in the real world (Kaushik & Walsh, 2019) and therefore is particularly pertinent in the current context: a substantial number of children starting school without the expected level of language skills for their age is a clear example of a real-world problem, and a variety of methods may be appropriate to achieve the end goal of improving children's language and associated achievement and well-being indicators.

Methodology: a Corpus-Based Approach

Studying language acquisition in young children can involve various methodologies, including observational and experimental approaches, and analysis of existing databases to derive new knowledge (e.g., age of acquisition data). Where research is exploratory in nature, as is the case here, the collection of observational data tends to be a precursor in order to develop hypotheses that can be tested under experimental conditions. One way of collecting observational data is to use a corpus-based approach. A corpus is a body of written text or transcribed speech that can be analysed from a linguistic perspective (Kennedy, 1998) and is usually machine-readable and representative of a particular language or form (McEnery et al., 2006). In a broad sense, corpus linguistics is any form of linguistic inquiry based on data derived from samples of language produced in genuine communicative situations (Stefanowitsch, 2020). Corpus linguistics is a rapidly developing field, and as such there is no 'magisterial perspective' (Kennedy, 1998, p.2) regarding a unified set of methods or accepted

framework, but approaches tend to utilise specialised software search tools to analyse large sets of data (containing machine-readable texts), producing both quantitative and qualitative data as required, and to be employed alongside other approaches where necessary.

Corpus data provides a rich source of naturally occurring information with high ecological validity (Ventura et al., 2021; Hoff, 2011), it is thus well suited to facilitate a data driven approach to address the current research question: investigating the linguistic nature of the vocabulary gap in pre-schoolers. By extracting and analysing frequency data from the spoken transcripts of children in an everyday setting, patterns can be identified and subsequently associated with particular groups of children who may need extra help to reach the required levels of language needed for school and beyond. Widely used systematic procedures for data collection, transcription and analysis mean that naturalistic data of this kind is now more amenable to scientific analysis than it has been in the past, and has therefore increased its scientific value (Lust & Blume, 2017).

An alternative approach for this exploratory analysis would have been to use language assessment data as opposed to corpus data, whereby the words children were able to express when prompted or via parental report could be used as units of analysis instead of the words they actually used during the naturalistic observation. Well-established tests such as the MacArthur-Bates CDI, PPVT or BAS are commonly and successfully utilised by researchers to draw conclusions about different aspects of children's language outcomes, including the influence of infant directed speech (Hartman et al., 2017) and age of acquisition research (Braginsky et al., 2016). For this particular project however, observational data was more appropriate, particularly as the data was secondary in nature. This approach provided rich raw data that could be analysed at a chosen level (in this case word level), rather than secondary assessment data, which is often provided as an overall score – as was the case with the British Ability Scales (Naming Vocabulary subtest) used by the Millennium Cohort Study. One overall vocabulary score would have been insufficient to address the first research objective, which required analysis of the patterns of individual word usage in order to ascertain which word types tended to be associated with large and small vocabularies. Using naturalistic corpus data also facilitated a context where there were no constraints on the words that children could use, as even the longest and most comprehensive language assessment only provides an index of children's knowledge, with high probability of a ceiling effect.

Using naturalistic corpus data is not without its limitations, however. A child's linguistic knowledge (in this case their productive vocabulary) can only be inferred from any observed linguistic behaviour (Ambridge & Rowland, 2013), and their behaviour in any given situation can be influenced by a range of factors. The range of toys and books available, for example, is likely to influence children's speech in different ways depending on their preferences, and does not always elicit more speech (Lust & Blume, 2016). How much parents interact with their children is also likely to influence the child's own speech, as is the time of day when the recording takes place (for example particular routines such as bathtimes and mealtimes may be more likely to elicit speech than others). Other contextual factors such as the presence of siblings or anything else which may provide auditory distractions (Beaman et al., 2014) are also likely to influence a child's speech production within any given time window.

There are also limitations associated with transcribing the data, in terms of the time-consuming nature of transcribing (Hoff, 2011), and ensuring accuracy. The latter can be in some way

mitigated by repeated checks for errors (as was done here) and is helped by the repetition of difficult-to-decipher utterances by parents (Moskowitz, 2000, as cited in Lust & Blume, 2016), or the use of gesture in the videos for clarification. For the current study, the snapshot of observed behaviour is relatively small (two x 10 minute videos), which has potential to lead to over or under-estimation of any given speech phenomena (Ambridge & Rowland, 2013), and yet the fact that the current study is concerned with all the speech which children produce within the given time window (as opposed to particular word combinations, for example), suggests this will be less of an issue. These limitations will be considered in more detail in the Chapter 6.

Ultimately, the research paradigm chosen accommodates a range of methodological approaches by primarily focusing on and addressing the stated research question. Research that details the discrepancy in language abilities between pre-schoolers comes from a variety of valuable data sources, including teacher surveys (Harley et al., 2018), EYFS assessment data (Law et al., 2017a), corpus data (Weisleder & Fernald, 2013) and experimental data (Fernald et al., 2013), therefore it is appropriate that any research aiming to pinpoint the exact linguistic nature of the discrepancy or gap also comes from a variety of sources – with the current project focusing on only one of those sources (corpus data).

Using data from different sources also provides a useful triangulation point, whereby multiple methods or data sources are used to develop a more comprehensive understanding of a phenomenon (Patton, 1999), allowing for greater levels of confidence in any findings where the same results are obtained using different methods or datasets (Rothbauer, 2008). This is particularly important when research is new and exploratory, or where there is contradictory evidence regarding a particular issue. Triangulation fits particularly well within the pragmatic research paradigm, which specifically allows for and even encourages a range of approaches to achieve the overall goal of achieving the research aims (Dewey, 1910). For this project, BAS (naming vocabulary subtest) scores from the Millennium Cohort Study (a large-scale study following the lives of around 19,000 children from birth through their entire childhood, with multiple data points) have provided a useful triangulation point with respect to which demographic variables tend to be most associated with language outcomes.

Data Collection: the Sample

Data was obtained via a legal data sharing agreement (see Appendix A) with LuCiD (The ESRC International Centre for Language & Communicative Development), and consisted of two individual ten-minute video files for each of 73 child participants aged 36 months old, one showing the child playing independently, and the other showing the child playing with their parent. Filming took place indoors within a medium-sized room, and the children had access to a range of age-appropriate toys and books. Present in the room for each session was the LuCiD researcher (operating the video camera), the parent, and occasionally a younger sibling. For the independent play session, the parent was still present in the room, but did not actively interact with the child.

The data was collected as part of LuCiD's longitudinal Language 0-5 Project, which followed 80 children over a period of 48 months, from 6 to 54 months, collecting a variety of naturalistic, experimental and standardised measure data. This was the largest and most recent collection of speech data for children of this age group known to the researcher at the time (data collected 2014-19). The specific data requested and obtained related to the 36-month data sweep, along with

additional demographic and standardised assessment data for both the initial 6-month and later 36-month data sweeps. This age group was particularly suitable for the current investigation, in that it was located at the lower end of the pre-school age range (3-5 years), an age where children are old enough to have a large enough vocabulary to sufficiently analyse, but also young enough to still benefit from a parent-led intervention at home, as intended for the intervention Action Research cycles in Phase Two of this thesis. Vocabulary measures at 3 years old have also been shown to be more valid predictors of later language then those taken at an earlier age (Dale & Kachergis, 2023).

Preparing the Data: the Transcription Process

The videos provided by LuCiD were transcribed using the CHAT transcription system and the CLAN (computerised language analysis) program. CLAN allows transcripts to be linked directly to the media files being transcribed, allowing for a speedier and more accurate process. The CHAT system is a series of conventions and principles applied to the process of creating transcripts so that they can be compared and used by different researchers. These conventions include using a main tier for recording the words spoken, and a dependent tier running underneath for any extra information about those words. Both systems are linked to CHILDES (Child Language Data Exchange System), an online databank of freely available transcripts of children's speech, and manuals supporting their use can be found linked from the website (MacWhinney, 2000).

Listening to each video recording in turn, 146 CHAT files were created, one for each video file provided (two per child participant). See Appendix B for CHAT files (example screenshots of first page, due to file format). Despite the current study focusing on the child's speech only, full transcripts were nonetheless created for thoroughness and potential future use (containing child, parent and researcher speech). To check for reliability, each recording and related transcription was checked two additional times after the first initial listening and transcription (by the same researcher) to reduce the risk of human error. In circumstances where a particular word could not be deciphered (due to young children's talk sometimes being unintelligible or background noise) further listening and watching of the video could usually shed light on the child's intended word. When the word could not be ascertained after the third time of listening and watching, it would be denoted as *XXX* and subsequently ignored by the programme. Babbling, denoted as e.g., *&gaga* was also ignored. Interjections and communicators such as *ahem*, *ugh* and *uh* were included in the form indicated in the co.cut file so that they would be recognised by the MOR programme (see p.56 of CHAT manual linked from CHILDES; Macwhinney, 2000).

Whilst most analysis was to be carried out using R, it should be noted that CLAN has considerable analytic capabilities of its own. The MOR programme, which has a similar user manual linked to the CHILDES databank (MacWhinney, 2000), is particularly notable, and was run on each of the 146 CHAT files. MOR provided morphological coding for every word transcribed by adding a dependent tier directly underneath the main tier. This coding included word category information (for example noun, verb, preposition etc), as well as allowing VOCD (vocabulary diversity) to be calculated for each child at a later stage.

The morphological coding provided by the MOR programme is based on a substantial vocabulary database, however due to differences between US and English spellings, it was necessary to cross-check the MOR word categories to ensure that UK spellings had been recognised, and where

they had not, to enter the information manually. This was done for *trapezoid* versus *trapezium*, and *color* versus *colour*. Morphological coding by the MOR programme was also altered manually where a 'childspeak' version of a word had not been recognised (e.g., *horsey* versus *horse*), and where there had been contextual errors (e.g., *orange* may have been categorised as an adjective by MOR, but further investigation showed that the context in which it was used was in fact a noun). This cross-checking was done after the CSV files had been created (see below).

Once the CHAT files were complete, they were then converted to CSV files ready for statistical analysis in R. This was done by using the ELAN programme to extract only the child's words from the main line of the transcripts and then creating a text file of all the words spoken by the child for each transcript. In order to maximise the words available for analysis (and because for this study, there was no interest in comparing the child's verbal behaviour when they were playing alone versus playing with an adult) the two text files for each child (created from the independent play session and the interactive play session with a parent) were combined within Excel, and then converted to a CSV ready for analysis using R. This process was repeated for the mor % line (the dependent tier) of the transcripts containing morphological information, so that in the end each child had two associated CSV files: one containing all the words they spoke over the two video play sessions, and the other containing all the morphological information associated with those words. These files were named, for example, Am26_words and Am26_wordcat, using the first two letters of the child's name and their participant number. It was important to retain this information in the file name so that the videos could be returned to in the case of any glitches. As already mentioned, the word and wordcat files were cross-checked to make sure that each word had been categorised correctly. Where there was ambiguity, the video was re-checked to ascertain the context of the word's use (e.g., the word *drink* can be utilised as a verb or a noun).

Linguistic Measures

In order to address Research Objective 1 regarding the linguistic nature of the vocabulary gap, a number of linguistic measures were chosen. These choices were based on the available evidence and there is a detailed review and rationale for each chosen variable in the previous chapter. Each variable focuses on word-level factors (such as a word's lexical category or phonological neighbourhood density) as opposed to those which would require analysis of the context in which the word was spoken. These latter factors, including the use of decontextualised language (Rowe, 2012) and the referential transparency of the exchange (Cartmill et al., 2013) have also been associated with larger vocabularies, but would not be directly relevant to answering the current research question about the specific word types that reside within the vocabulary gap, as whether or not something could be considered decontextualised or not would vary depending on the situation, and referential transparency is about the non-verbal features of speech. As such, any findings from the current study should be interpreted alongside such evidence.

Word Category

Morphological coding by the CLAN program provided an initial categorisation of words, and further categorisation was done using R code to create fewer and broader categories, somewhat similar to Bates et al. (1994) and Braginsky et al. (2019), who categorised their words into three groups: nouns, predicates and function words, where predicates included verbs and adjectives in the former

study, and verbs, adjectives and adverbs in the latter. For the current study, four groups were used as follows: all words labelled as nouns or proper nouns were classed as 'noun', all words labelled as adjectives were classed as 'adjective', all words labelled as verbs, participles, auxiliary verbs, modal verbs or copula verbs were classed as 'verb', and all other words (provided they were not associated with any of the other main categories: noun, adjective and verb) were classed as 'function'. In alignment with the pragmatic approach of this research, the reason for grouping the words in this way, and in particular the splitting out of the predicate category into verb and adjective, is the practically useful result it was hoped would be achieved: the terms 'noun,' 'verb,' and 'adjective' (as opposed to 'predicate') are commonly used and understood by education professionals and parents, and to see the individual role they play was therefore important. The term 'adverb' is also commonly used and understood, however often in its simplest role (as a modifier of another verb, for example *quickly* or *happily*), and many more complex adverb types (for example those related to time) probably require a greater degree of conceptual understanding that go beyond the current age group of interest (Kowalsky, 2008), and thus adverbs were included in the function category as opposed to a stand-alone category. Raw counts were used, rather than proportions (similar to Rowe, 2012) to avoid issues of using proportion data that adds up to one, and therefore is not linearly independent.

Word Frequency

Word frequency refers to a measure of how commonly the word appears in the language. In this case, frequency scores for each word were calculated by utilising the British National Corpus (BNC), a collection of 100 million words from a wide variety of written and spoken language from the late twentieth century (a subset from the spoken BNC data was used, containing 67,521 words, as was used by Jones et al., 2021). This included speech from unscripted informal conversations and a range of other more formal contexts, including classrooms, business meetings and sports commentaries, and was balanced across different age-groups and backgrounds. Possible scores ranged from 1–409,650, with high scores indicating a word with higher frequency (and therefore appears more commonly across the language).

Word Length

Word length can be measured using the number of letters, phonemes or syllables, all of which are correlated (Lewis & Frank, 2016). For this project, number of phonemes are used as a proxy for length, similar to the English Lexicon Project (Balota et al., 2007; Jones & Brandt, 2019). The number of phonemes in each word was calculated by utilising the Carnegie Mellon University (CMU) Pronouncing Dictionary, an online and open-source machine-readable pronunciation dictionary for North American English that contains over 134,000 words and their pronunciations.

Pronunciation Difficulty

There is a lack of consistency in existing literature concerning how to operationalise this measure. Braginsky et al. (2019) used number of phonemes as a proxy for pronunciation difficulty, but in this case, that would not differentiate from word length (above), and hence number of syllables was chosen as an alternative proxy. Again, the number of syllables was calculated by utilising the Carnegie Mellon University (CMU) Pronouncing Dictionary (as above).

Rareness

A word's rareness was identified by utilising the Dale-Chall word list (Chall & Dale, 1995), that contains approximately three thousand familiar words that are known (as judged by teachers) in reading by at least 80% of Grade 4 children in the USA, as used by Rowe (2012) and Weizman and Snow (2001). Any word not on the list was given a 'rare' rating, whereas words that were on the list were considered 'not rare.' Manual checks were completed after the look up table code had been run to ensure consistency, and as a result some words were transferred from rare to not rare in a two-part process, which first involved a subjective consideration of whether a word could sensibly considered rare or not on a case-by-case basis, and then, if a decision was made to remove the word from the rare list, then a reason code was assigned to this decision - for example the word is a proper noun, present on CDI forms (indicating commonality), or a common interjector or exclamation. Despite involving an element of subjectivity, from a pragmatic perspective this was a necessary intervention to avoid the inappropriate rare labelling of some words arising due to, for example, the discrepancies between US and English spellings or language conventions (e.g., *mummy*, *teddy* and *granny* were all initially labelled as rare), and thus resulting in a skewed picture of the number of rare words children produced. A list of exceptions was therefore created (see Appendix C), containing 101 exceptions alongside their reasons codes as described above. A similar approach was taken in Weizman and Snow (2001), where, for example, names of family members, forms of address (e.g., *Mommy*, *Honey*) and popular child culture terms (e.g., *Power Rangers*) were also removed from the rare category on a case-by-case basis. This list of exceptions was then used as an additional look up table within the overall coding script. As with word categories, raw counts were used, rather than proportions (see explanation above).

Concreteness

Concreteness refers to a word's tangibility versus its abstractness. A concreteness value was calculated for each word by utilising the Brysbaert et al. (2014) database (also used by Braginsky et al., 2019, and Jones & Brandt, 2019). The database contains 37,058 words rated for their concreteness. Scores ranged between 1-5, with high scores indicating more tangible and concrete words, and low scores indicating less tangible and more abstract words.

Phonological Neighbourhood Density (PND)

PND refers to the number of words that can be formed by addition, substitution or elimination of a single phoneme in target word (for example 'cat' has a PND of 32). PND values were calculated for each word by utilising the English Lexicon Project (Balota et al., 2007, Jones & Brandt, 2019) a database of 40,481 words and their lexical characteristics. Possible scores ranged between 0-49, with a higher score indicating a word with a greater number of phonological neighbours.

Phonotactic Probability (PP)

Phonotactics is the sequential arrangement of phonological units that are possible in a language (Crystal, 1987), in this case English. Some arrangements are possible, for example /str-/ and some are not, for example /stb-/. PP refers to the likelihood of a particular sound sequence within a word occurring again within a particular language sample, for example the sound sequence /-est/ at

the end of a word is more common in English than the sound sequence /-xt/. PP values were calculated for each word by utilising the Irvine Phonotactic Online Dictionary (IPHOD, Vaden et al., 2009), a large collection of around 54,000 English words and pseudowords developed for research on speech perception and production and containing information based on measures related to phonemes. The measure of PP used for each word was 'unsLBPAV' that refers to the log (base 10) statistic for unstressed word average biphoneme probability (relative frequencies for ordered phoneme pairs). Possible scores ranged between 0-0.027, with a higher score indicating a word with a more common sound sequence.

Demographic Measures

Research Objective 2 was to investigate the most salient demographic risk factors associated with smaller vocabularies, with a view to eventually designing and implementing an intervention targeted at this group. Included within the legal data sharing agreement with LuCiD were numerous associated demographic and standardised assessment data for each participant from both the initial 6-month and later 36-month data sweep. Nine variables from the LuCiD dataset (necessarily limited in scope by what was available) were chosen to be included in the analysis: gender, number of siblings, maternal age, paternal age, maternal education, paternal education, income, paternal caregiving responsibilities and dummy usage. These choices were based on existing evidence considered and reviewed in Chapter 2 which suggested each variable to be an important factor influencing language development. The purpose of completing additional analysis on the MCS dataset was triangulation; to corroborate or otherwise the story being told by the LuCiD data. Eight variables from the MCS dataset were therefore chosen, with seven of these variables matched with the LuCiD variables: gender, income, no. of siblings, maternal age, paternal age, maternal education and paternal education. An eighth variable in the form of employment status was also included as a supplementary point of interest - whilst not intended as a triangulation point (there was no comparable variable included for the LuCiD dataset) - employment status has been linked to children's overall achievement (Jenson, 2023) and literacy outcomes (Siddiqi et al., 2007), and therefore it was expected that unemployment, particularly when applicable to both parents, would have a negative influence on children's language development, either through loss of income (Jenson, 2023) or some other pathway (Brand, 2015). Statistics pertaining to employment status within the MCS sample would also be useful for descriptive purposes. The MCS had no comparable data on paternal caregiving responsibilities or dummy usage.

LuCiD Dataset

Gender. Data collected by LuCiD allowed for two categories: male or female. All participants were therefore coded as either 1 for male, or 0 for female.

Number of Siblings. The number of siblings the child participant has at the time of the video recording (when the child was 36 months old). Answers ranged from 0 to 4 or more, with answers of 4 or more recoded as 4. Scores therefore ranged from 0-4, with higher scores indicating a greater number of children in the household.

Maternal Age. The mother's age when the child participant was 6 months old. Answers were categorised into age bands, and coded as 0 for under 21 years, 1 for 21-25 years, 2 for 26-30 years, 3

for 31-35 years, and 4 for over 35 years, therefore scores ranged from 0-4, with higher scores indicating an older age.

Paternal Age. The father's age when the child participant was 6 months old. Answers were categorised into age bands, and coded as 0 for under 21 years, 1 for 21-25 years, 2 for 26-30 years, 3 for 31-35 years, and 4 for over 35 years, therefore scores ranged from 0-4, with higher scores indicating an older age.

Maternal Education. The mother's highest level of formal education received when the child participant was 6 months old. Answers were coded as 0 for no qualifications, 1 for GCSE or equivalent, 2 for A-level or equivalent, 3 for degree or equivalent, and 4 for postgraduate qualification or equivalent, therefore scores ranged from 0-4, with higher scores indicating a higher level of education.

Paternal Education. The father's highest level of formal education received when the child participant was 6 months old. Answers were coded as 0 for no qualifications, 1 for GCSE or equivalent, 2 for A-level or equivalent, 3 for degree or equivalent, and 4 for postgraduate qualification or equivalent, therefore scores ranged from 0-4, with higher scores indicating a higher level of education.

Income. The level of annual income household received at the time of the video recording (when the child was 36 months old). Answers were categorised into income bands, and coded as 0 for NA, 1 for £14,000 - £24,000, 2 for £24,000 - £42,000 and 3 for over £42,000, therefore scores ranged from 0-3, with higher scores indicating a higher income.

Paternal Caregiving Responsibilities. The father's caring responsibilities for the child participant at the time of the video recording (when the child was 36 months old). Coded as either 1 for yes, or 0 for no.

Dummy Usage. The child participant's use of dummies, historical and present, at the time of the video recording (when the child was 36 months old). Answers were coded as 0 for "no, never", 1 for "yes, but stopped before 30 months", and 2 for "yes, still do," therefore scores ranged from 1-3, with higher scores indicating a higher level of dummy usage.

MCS Dataset

The MCS dataset contained 7652 participants after missing data was removed, as detailed in the sections below.

Gender. All participants were coded as either 1 for male, or 2 for female, or -1 for NA. Participants with -1 were removed.

Number of Children in the Household. The total number of children (including the child participant) in the household at the 36-month data point. Answers ranged from 1 to 13, with higher scores indicating more children in the household. Answers given as -1 or -2, equating to not known or NA, were removed.

Employment Status. The employment status of both adults in the house. Participants with scores of -9, -8, -1 were removed (equating to refusal, NA and don't know) and the remaining answers were

coded as 1 for 'both in work', 2 for 'main in work, partner not' and 'partner in work, main not', and 3 for 'both not in work.' Therefore, those scoring 1 were considered to have a higher employment status than those scoring 2 or 3. Answers scoring between 5-11 were removed due to the complex nature of deciding where they should sit within the ordinal nature of 'employment status' (e.g., 'main status unknown, partner in work'). Scores therefore ranged from 1-3, with higher scores indicating lower employment status.

Income. The level of income received by the whole household. Scores of -9, -8 or -1 were removed (equating to refusal, NA and don't know). Answers were categorised into income bands, and coded as 1 for £0 - £3,300, 2 for £3,300 - £11,000, 3 for £11,000 - £22,000, 4 for £22,000 - £33,000, 5 for £33,000 - £55,000, and 6 for £55,000 and over, therefore scores ranged from 1-6, with higher scores indicating a higher income.

Maternal Age. The mother's age at the birth of the child participant. Answers of -1 and -2 were removed (indicating not known or NA). Answers were categorised into age bands, and coded as 1 for 11-19, 2 for 20-29 years, 3 for 30-39 years, 4 for 40 years and over, therefore scores ranged from 1-4, with higher scores indicating an older age.

Paternal Age. The father's age at the birth of the child participant. Answers of -1 and -2 were removed (indicating not known or NA). Answers were categorised into age bands, and coded as 1 for 11-19, 2 for 20-29 years, 3 for 30-39 years, 4 for 40 years and over, therefore scores ranged from 1-4, with higher scores indicating an older age.

Maternal Education. The mother's highest level of formal education received at the time of the 36-month data point. Scores of -9, -8 or -1 were removed (equating to refusal, NA and don't know), as were scores of 95 and 96 (indicating 'overseas qualification only' or 'none of these'). Answers were coded as 1 for NVQ level 1, 2 for NVQ level 2, 3 for NVQ level 3, 4 for NVQ level 4, and 5 for NVQ level 5, therefore scores ranged from 1-5, with higher scores indicating a higher level of education.

Paternal Education. The father's highest level of formal education received at the time of the 36-month data point. Scores of -9, -8 or -1 were removed (equating to refusal, NA and don't know), as were scores of 95 and 96 (indicating 'overseas qualification only' or 'none of these'). Answers were coded as 1 for NVQ level 1, 2 for NVQ level 2, 3 for NVQ level 3, 4 for NVQ level 4, and 5 for NVQ level 5, therefore scores ranged from 1-5, with higher scores indicating a higher level of education.

Measures of Vocabulary

Lexical diversity (a measure of children's vocabulary size) was operationalised by two different measures, with each being included as a dependent variable in a separate model. The first was vocabulary diversity (VOCD), and the second number of different words (NDW). VOCD controls for sample size by accounting for the fact that samples with larger token sizes tend to have lower type-token ratios, as each time a new word is added, this increases the token count, but only increases the type count if the word has not been used before, hence TTR will inevitably fall as sample size increases (Durán et al., 2004). Calculated using CLAN, the VOCD measure represents the probability of new vocabulary being introduced as sample size increases, with the overall score derived from the type-token ratio versus the token curve calculated from data for the transcript as a whole (Jones & Rowland,

2017). It is a broad interpretation of diversity in that it is about more than just vocabulary range and size, but also about the way vocabulary is deployed beyond a word's initial use (Malvern et al., 2004).

The second measure of lexical diversity used was the total number of different words used by the child over the combined ten-minute videos (NDW). Whilst NDW is clearly affected by the size of the speech sample, in this case the samples of speech were all standardised on time (each child was recorded for 10 minutes alone and 10 minutes with a parent), thus allowing NDW to be utilised as both a simple and standardised measure of children's vocabulary size. Including both VOCD and NDW in the models was a considered and deliberate decision, not just because two differently calculated measures of the same thing provide a useful triangulation, but also because whilst VOCD is likely to be the most commonly used measure of lexical diversity, it is not universally used by researchers. For the results of this study to be directly comparable to a range of studies it was therefore sensible to include both measures, whilst accepting that the NDW measure can only be directly compared where the same time window has been used. Further consideration will be given to this decision in the discussion chapter.

Two additional measures of vocabulary were provided by LuCiD in the form of MacArthur CDI (Communicative Development Inventory) Lincoln CDI and MacArthur CDI-III scores, and these were used as useful comparisons and descriptors of the dataset but were not included in any of the models. For the MCS dataset, scores were provided for the BAS (British Ability Scales) Naming Vocabulary subtest, which focuses mainly on children's knowledge of nouns via their recognition of pictures.

VOCD (Vocabulary Diversity)

A measure of vocabulary diversity that accounts for sample size and is calculated from each child's transcript using CLAN (Duran et al., 2004). Scores ranged from 19.58 to 61.54, with higher scores indicating better vocabulary diversity. VOCD has been shown to correlate strongly (r=.95) with other measures of lexical diversity (deBoer, 2014).

NDW (Number of Different Words)

A measure of vocabulary diversity calculated by totalling the number of different words used by each child during the 20-minute video session. Scores ranged from 40 to 227, with higher scores indicating a greater number of different words used.

MacArthur Communicative Development Inventory: Level III (CDI-III)

The CDI-III (Fenson et al., 2007) is a parent report measure of productive language out of a total score of 100, and aimed at 30 to 37-month-olds. Scores ranged from 26 to 99, with higher scores indicating a greater number of words produced by the child. Internal consistency for the 100-item checklist is indicated by an alpha value of α=.98 (Dale & Kachergis, 2023, p.131), with concurrent validity estimates with measures (e.g., Bayley Scales of Infant and Toddler Development) moderate or above (Dale & Kachergis, 2023, p.132).

Lincoln Toddler Communicative Development Inventory (TCDI)

The Lincoln TCDI (Meints & Fletcher, 2011) is a parent report measure, including productive language out of a total score of 689, specifically adapted for UK use, and aimed at 16 to 30-month-olds. Scores ranged from 24 to 689, with higher scores indicating a greater number of words produced by the child. Psychometric properties on this instrument have yet to be published by the authors (Just, 2018).

British Ability Scales II (BAS-II), Naming Vocabulary Sub-test (Standardised Score)

The BAS-II (Elliott, 1996), naming vocabulary sub-test is a measure of expressive language assessed by asking children to name pictures of objects in a booklet. It is aimed at children aged between 2 years and 6 months to 5 years and 11 months, and depends on noun knowledge and picture recognition abilities. The BAS-II overall (consisting of a battery of 12 core sub-tests) has demonstrated both good construct validity as a measure of cognitive ability (Elliott et al., 1997), and high test-retest reliability (Elliott et al., 1997). T- scores ranged from -8 to 80, with higher scores indicating a greater number of words known and used by the child, and negative scores indicating the test was finished early or not carried out. Participants with negative scores were removed.

Methods of Statistical Analysis

1. All transcripts and data tables were uploaded to R in the form of excel or CSV documents. Where the existing data format was different (for example the transcripts were in CHAT format), then they were first converted into one of the accepted formats before uploading, as explained above.

2. For VOCD scores to be calculated for each child, their two associated transcripts (independent play and interactive play with a parent) were combined into one, and the VOCD command was run in CLAN on each combined transcript, hence producing an overall VOCD score for each child.

2. Using a looping code, all transcripts were compared and scored against each linguistic variable look up table (e.g., the BNC), providing a summary table of linguistic profiles for each participant, with 'linguistic profile' referring to the mean values or (in the case of rare words and the different word categories) raw counts for each of the linguistic measures described above.

3. LuCiD demographic data, CDI language scores and VOCD scores were combined from various data sources and where necessary, descriptive variables were recoded with numerical values. This information was then added to the linguistic profiles, creating a summary table of all demographic and linguistic information for all participants.

4. MCS demographic and language score data was also combined and where necessary, descriptive variables were recoded with numerical values and participants with scores of NA removed. An overall summary table was created including a BAS language score along with the chosen demographic variable scores for each child.

5. Two stepwise regressions were carried out involving the LuCiD linguistic and demographic variables, with Model 1 using VOCD as the dependent variable, and Model 2 using NDW as the

dependent variable. A stepwise model was chosen (as opposed to a multiple regression with a limited number of chosen variables) as this allowed a large number of variables to be included – all of which had been shown to be important for children's language outcomes - and for the most important predictors to be selected methodically by the process itself.

6. For the MCS data, where there were both less predictor variables and vastly more participants, all predictors were included in a multiple regression model (Model 3), with BAS scores as the dependent variable.

7. As a consequence of the results obtained in 5., further analysis was then carried out to ascertain whether or not those children who used a greater number of verb types (hence were more 'verbally diverse') throughout their transcript tended to have a higher lexical diversity in the utterances that contained one of these verbs, as compared to a less verbally diverse child using the same verbs.

8. To facilitate this analysis, the ten most frequent verbs used by the children as a whole cohort were identified using R, and then every utterance in every transcript featuring one of these verbs was identified using CLAN, and new 'subset' transcripts created for every child, including only those utterances that contain one of the frequent verbs identified. Two measures of vocabulary diversity (VOCD and NDW) were calculated for each subset, with 8 participants removed due to insufficient data. Regression analysis was then carried out to ascertain whether verb diversity (as measured by the number of verbs children used overall) predicted lexical diversity in the subset of common verbs. The code scripts are included in Appendices D-H.

Chapter 5: Results (Phase One)

Descriptive Statistics

Descriptive statistics can be found be in Table 1 (LuCiD Linguistic Variables), Table 2 (Descriptive Statistics and Correlations for LuCiD Vocabulary Measures) , Table 3 (Descriptive statistics and Norm Comparisons for LuCiD Vocabulary Measures), Table 4 (LuCiD Demographic Variables) and Table 5 (MCS Variables).

Table 1

Descriptive Statistics for LuCiD Linguistic Variables (Participant Age = 36 months, n = 73)

Variable	Mean	SD	Coefficient of Variance	Min.	Max.
Mean Syllables	1.29	0.05	0.04	1.19	1.4
Mean Phonemes	3.35	0.15	0.04	3.07	3.71
Mean Concreteness	3.08	0.14	0.05	13.55	18.32
Mean PND	15.58	0.92	0.06	13.55	18.32
Mean PP	0.0035	0.00029	0.08	0.0027	0.0042
Mean Frequency	34088.82	6286.24	0.18	22245.07	57124.62
No. of Verbs	33.03	11.15	0.34	7	62
No. of Adjectives	6.23	3.45	0.55	0	15
No. of Rare Words	4.53	3.19	0.7	0	16

Mean Syllables: mean syllable count for words spoken by each participant, Mean Phonemes: mean syllable count for words spoken by each participant, Mean Concreteness: mean concrete rating for words spoken by each participant, Mean PND: mean Phonological Neighbourhood Density score for words spoken by each participant, Mean PP: mean Phonotactic Probability score for words spoken by each participant, Mean Frequency: mean frequency rating for words spoken by each participant, No. of Verbs: number of verbs spoken by each participant, No. of Adjectives: number of adjectives spoken by each participant, No. of Rare Words: number of rare words spoken by each participant.

*Only variables chosen by the stepwise regression are included in the descriptives table, hence no. of nouns and no. of functions have been removed. NDW was measured by 2 different programmes (CLAN and R); only the R version is included in the table for simplicity and due to both measurements producing same results in regression analysis.

Coefficient of variance values of less than 1 for all variables indicate low levels of variability within the data. There is limited reliable data available for comparison to ascertain whether the listed linguistic measures are representative of the population. It would thus be informative to replicate with an independent sample.

Table 2

Descriptive Statistics and Correlations for LuCiD Vocabulary Measures (Participant Age = 30 - 36 months)

Variable	n	Mean	SD	Coefficient of Variance	VOCD	NDW	CDI III	CDI Lincoln
VOCD	73	36.77	9.51	0.26	-			

NDW	73	124.85	36.8	0.29	.43***	-		
CDI III	69	77.32	16.52	0.21	-.01	.24**	-	
CDI Lincoln	60	581.63	121.86	0.21	.09	.26	.71***	-

VOCD: vocabulary diversity, NDW: number of different words as calculated by R, CDI III: productive vocabulary measured at 36 months, CDI Lincoln: productive vocabulary measured at 30 months.

A mean VOCD of 36.77 indicates the sample was less lexically diverse overall than the sample in Duran et al (2004), where the mean VOCD was 47.83 for children aged 36 months. There is limited reliable data available for comparison regarding the NDW measure (which can only be directly compared with studies using the same time window), it would thus be informative to replicate with an independent sample. A mean CDI-III score of 77.32 also indicates children in the sample had smaller productive vocabularies when normed against 131 US children in the MacArthur-Bates CDI User's Guide and Technical Manual (Marchman et al, 2023, p.363-364, p.372) where the 50th percentile score at 36-37 months was 83.5 (see Table 3). On the other hand, the mean score for the Lincoln-CDI administered at 30 months was 581.63, which is somewhat higher than the 50th percentile score indicated by the same source of norming data (Marchman et al, 2023, p.289-209, p.322). Implications for these different findings will be discussed in Chapter 6.

Zero-order correlations between the vocabulary measures indicate moderate-strong significant correlations between the two measures of vocabulary based on children's speech samples (VOCD and NDW), and between the two parent report measures (CDI-III and CDI-Lincoln), as well as a weak significant correlation between CDI-III and NDW. These results indicate a degree of discrepancy between the vocabulary measures used, the implications of which will again be discussed in Chapter 6.

Table 3

Descriptive Statistics and Norm Comparisons for LuCiD Vocabulary Measures (Participant Age = 30 - 36 months)

Variable	n	Min.	Max.	Norm referenced percentile scores			% of sample			
				10th percentile	50th percentile	90th percentile	< 10th percentile	≤ 50th percentile	> 50th percentile	> 90th percentile
VOCD	73	19.58	61.54	30 [1]	47.83	65 [1]	28.8	89	10.96	0
NDW	73	40	227	-	-	-	-	-	-	-
CDI III	69	26	99	52.5	83.5	98	8.7	60.87	39.13	2.9
CDI Lincoln	60	24	689	223.5	509	659	1.67	18.33	81.67	26.67

VOCD: vocabulary diversity, NDW: number of different words as calculated by R, CDI III: productive vocabulary measured at 36 months, CDI Lincoln: productive vocabulary measured at 30 months. Sources of norm referenced percentile scores referenced below.

1. This is an estimated figure, see Duran et al (2004, p.230)

Table 3 provides indications of range for each measure of vocabulary, both within the sample (Min. and Max. scores), and then in comparison to norm-referenced percentile scores. Indications of

range are important for the particular research aims being addressed here, which are concerned with the differences between children with larger and smaller vocabularies. These comparisons indicate that the vast majority of VOCD scores were distributed below the 50th percentile (Duran et al., 2004, p.230), with 28.8% below the 10th percentile, indicating a vocabulary spectrum which was skewed towards the lower end. CDI-III scores appear to be more evenly distributed when compared to norm referenced percentile scores (Marchman et al., 2023), although still concentrated below the 50th percentile. CDI-Lincoln scores, on the other hand, indicate a sample in which the vast majority (81.67%) of children are scoring above the 50th percentile, with 26.67% scoring above the 90th percentile. The implications of these observations will be discussed in Chapter 6, alongside any limitations associated with the norming data used to make comparisons here.

Table 4

Descriptive Statistics for LuCiD Demographic Variables (Participant Age = 6 months/36 months, n = 73)

	Absolute & Relative Frequencies (%)				
	Boys	Girls			
Gender	35	38			
%	48	52			
	£0	£14-24,000	£24-42,000	Over £42,000	
Income	1	3	20	49	
%	1.5	4	27.5	67	
	Under 21	21-25	26-30	31-35	Over 35
Mum Age	1	3	18	30	21
%	1	4	25	41	29
Dad Age	0	1	13	31	28
%	0	1.25	18	42.5	38.25
	None	GCSE/equiv.	A-Level/ equiv.	Degree/equiv.	Postgrad/equiv.
Mum Education	0	4	7	26	36
%	0	5	10	36	49
Dad Education	1	10	14	34	14
%	1	14	19	47	19

	Never used	Stopped pre-30 months	Stopped post 30 months	Still using
Dummy usage	37	14	3	19
%	51	19	4	26

	Yes	No
Dad caregiving	57	16
%	78	22

	Mean	Standard Deviation	Coefficient of Variance	Min.	Max.
No. of Siblings	0.77	0.99	1.29	0	4

No. of Siblings: number of siblings when participant aged 36 months, Gender: as determined at birth, Income: income measured when participant aged 36 months, Mum Age: mum's age in years when participant aged 6 months, Dad Age: dad's age in years when participant aged 6 months, Mum Education: mum's highest level of education when participant aged 6 months, Dad Education: dad's highest level of education when participant aged 6 months, Dummy Usage: participant's dummy usage up to 36 months of age, Dad Caregiving: whether or not dad had any caregiving responsibilities when participant aged 36 months.

The mean number of siblings for the LuCiD sample was 0.77, which when including the participant child, is very similar to the average total fertility rate of 1.71 in 2017, which can be used as a proxy for average number of children per household (Office for National Statistics, 2022). A coefficient of variance above 1 however shows that there is considerable variability in the data here. The proportion of girls and boys in the sample were 49% and 51% respectively, which again is broadly reflective of the UK population in general (Office for National Statistics, 2021). Relative frequencies indicate that most households in the sample earned over £42,000, which is considerably more than the UK median income of £28,358 at that time, (using equivalised figures deflated to 2017/18 prices, Office for National Statistics, 2018). Both women and men were mostly aged between 31-35 years when their child was 6 months old, although this was closely followed by those aged over 35. Again, this is reflective of the UK population as a whole, with the average age of parents (overall, not first-time) in the UK in 2017 being 30.5 years (mothers) and 33.4 years (fathers) (Office for National Statistics, 2017). Most parents in the sample were educated to postgraduate level (49% of females) or degree level (47% of males). In 2017, the UK population was reported to have 41% of adults educated to level 4 or above (StatsWales, 2023), with level 4 being equivalent to year 1 of a degree course. This indicates that the LuCiD sample parents held considerably higher educational qualifications, on average, than the general population.

Most fathers (78%) in the sample had caregiving responsibilities for their child, which seems to be a higher proportion than the general population where it is estimated that 6% of fathers regularly look after their children, with an additional 18% sharing childcare responsibilities equally with their partner (Poulter, 2010). The lack of additional information available however (from both LuCiD and the comparison source: Poulter, 2010) make direct comparisons difficult, as it is impossible to quantify the exact amount of childcare being carried out by fathers. Dummy usage was split between those who had never used one and those that had (even if they had subsequently given up). Dummy usage in the UK population is estimated to be around 64% (NHS Infant Feeding Survey, 2010), suggesting that the LuCiD sample had a slightly lower incidence of dummy usage than the population in general.

In summary, the LuCiD sample differs on a number of key measures when compared to the population in general, and therefore cannot be said to be completely representative. Whilst the number of siblings per child and parental age were similar to average, household incomes, parental qualifications and levels of paternal caregiving were higher than average, and dummy usage lower (although the latter two measures are difficult to quantify, as discussed).

Table 5

Descriptive Statistics for MCS variables (Participant Age = 36 months, n = 7652)

		Absolute & Relative Frequencies (%)				
		Boys	Girls			
Gender		3877	3775			
	%	51	49			
		£0-3300	£3300-11,000	£11,000-22,000	£22,000-33,000	£33,000-55,000
Income		141	390	1984	2236	2099
	%	2	5	26	29	27.5
		11-19	20-29	30-39	40+	
Mum Age		225	3175	4067	185	
	%	3	41.5	53	2.5	
Dad Age		80	2243	4619	710	
	%	1	29.5	60.5	9	
		NVQ Level1	NVQ Level 2	NVQ Level 3	NVQ Level 4	NVQ Level 5
Mum Education		463	2246	1331	3144	468
	%	6	29.5	17.5	41	6
Dad Education		509	2301	1348	2862	632
	%	7	30	18	37	8
		2 adults in work	1 adult in work	0 adults in work		
Employment Status		4874	2560	218		
	%	64	33	3		

	Mean	SD	Coefficient of Variance	Min.	Max.
BAS T Score	52.29	10.53	0.2	20	80
No. of Chn in House	2.11	0.9	0.43	1	13

BAS T Score: naming vocabulary t score, No. of Chn in House: number of children in household including participant and siblings, when participant aged 36 months, Gender: as determined at birth, Income: income measured when participant aged 36 months, Mum Age: mum's age in years at participant's birth, Dad Age: dad's age in years at participant's birth, Mum Education: mum's highest level of education when participant aged 36 months, Dad Education: dad's highest level of education when participant aged 36 months, Employment Status: combined employment status of parents/parent + partner when participant aged 36 months.

The mean BAS T (standardised) score for the naming vocabulary subtest was 52.29, which is very similar to the norming sample mean of 50, the norming sample being a group of 1689 children selected to be representative of the UK population in 1995 (Connelly, 2013). The mean no. of children in the household was 2.11, which is more than both the current average of 1.49 (Office for National Statistics, 2022) and the 2004 average total fertility rate of 1.77 (Office for National Statistics, 2022). These two taken together suggest that the children in the MCS sample were of average vocabulary skill, but had a higher-than-average number of siblings. The coefficient of variance for both variables is below 1, reflecting low levels of variability within the data.

The proportion of girls and boys in the sample were 49% and 51% respectively, which is reflective of the UK population overall (Office for National Statistics, 2021). Relative frequencies indicate that most households in the sample earned £22-33,000, which is comparable to the UK median household disposable income of £25,651 in 2004 (using equivalised figures deflated to 2017/18 prices) that will necessarily be slightly less than gross salary figures (Office for National Statistics, 2018), although a similar proportion indicated earnings either side of the median income bracket: £11-22,000 and £33-55,000. Both women and men were mostly aged between 30-39 years at the birth of their child, followed by those aged 20-29 years. Again, this is broadly reflective of the UK population at the time, with the average age of parents in the UK in 2004 being 29 years (mothers) and 32.1 years (fathers) (Office for National Statistics, 2017). The most common qualification level amongst parents in the sample (41% of mothers and 37% of fathers) was NVQ level 4, with the next most common highest qualification being NVQ level 2 (GCSEs). This indicates higher than average education levels among the cohort, as four years later in 2008, only 30% of working age adults in the UK had been educated to level 4 or above (StatsWales, 2023), with level 4 being equivalent to year 1 of a degree course. The percentage of parents who were both in work for the MCS sample is 64%, with another 33% of households with one parent working. Office for National Statistics data indicates that in 2004, 91% of men with dependent children were in employment, and 67% of women. This shows that both the MCS sample and the population in general have similarly low levels of whole household unemployment. In summary, the MCS cohort were broadly representative of the population in general at that time. The data indicated average values for household incomes, maternal and paternal ages, child vocabulary skills and employment levels. On the other hand, MCS families had a higher-than-average number of children in their household, and their qualification levels were above average.

Main Analysis

Model 1: LuCiD Variables with VOCD as Outcome Measure

In order to determine the relative influence of each of the variables listed above (see Tables 1 and 4) on children's vocabulary as per the aims of this research, a stepwise linear regression was performed incorporating all the LuCiD linguistic and demographic variables, with VOCD as the dependent variable. A stepwise model was chosen (as opposed to a multiple regression with a limited number of chosen variables) as this allowed a greater number of variables to be included – all of which had been shown to be important for children's language outcomes - and for the most important predictors to be selected methodically by the process itself. No. of Verbs, No. of Adjectives, Dad Age, Dad Education and Dummy Usage emerged as the predictor variables to include in the model, the results of which are shown below in Table 6.

Table 6

Summary of Regression Analysis for Predicting VOCD

	B	SE(B)	β	t	Sig.	
No. of Verbs	0.33	0.09	0.39	3.53	0.001	***
No. of Adjectives	0.51	0.31	0.19	1.67	0.101	
Dad Age	2.66	1.2	0.22	2.21	0.03	*
Dad Education	1.93	0.94	0.2	2.05	0.045	*
Dummy Usage	-1.31	0.73	-0.18	-1.8	0.076	.

No. of Verbs: number of verbs spoken by each participant, No. of Adjectives: number of adjectives spoken by each participant, Dad Age: dad's age in years when participant aged 6 months, Dad Education: dad's highest level of education when participant aged 6 months, Dummy Usage: participant's dummy usage up to 36 months of age.

*. p = 0.05-0.1 * p = 0.01-0.05 ** p=0.001-0.01 *** p=0-0.001*

The predictor variables were shown to explain 37% of variation in VOCD (multiple $R^2 = 0.37$), with the overall model being deemed as significant, $F(5,67) = 8.01$, $p < .001$.

Table 6 shows that No. of Verbs ($B = 0.33$, $t(67) = 3.53$, $p = .001$), Dad Age ($B = 2.66$, $t(67) = 2.21$, $p = .03$), and Dad Education ($B = 1.93$, $t(67) = 2.05$, $p = .045$) were significant predictors of VOCD, with dummy usage being of borderline significance ($B = -1.31$, $t(67) = -1.8$, $p = .076$): as the number of verbs, dad's age and education level increased, so did a child's VOCD (vocabulary diversity); greater dummy usage was associated with lower VOCD.

The standardised coefficients show that a one standard deviation rise in No. of Verbs predicted a 0.39 standard deviation rise in VOCD, a one standard deviation rise in Dad Age predicted a 0.22 standard deviation rise in VOCD, and a one standard deviation rise in Dad Education predicted a 0.2 standard deviation rise in VOCD. No. of Verbs was therefore the strongest predictor of VOCD in children, with a standardised coefficient of 0.39 indicating a large effect size (Cohen, 1988).

In an attempt to illustrate the real-world applicability of these results, it is important to consider and interpret the unstandardised coefficients as well as the standardised coefficients (Baguley, 2009). The former indicates that with each extra verb spoken, VOCD increases by 0.33 (range 20-62). As dad's age moves up into the next 5-year age band, VOCD increases by 2.63, and as his

education moves up into the next category, VOCD increases by 1.93. An increase in dummy usage by 1 unit (towards more usage) is associated with a decrease in VOCD by 1.31.

The core assumptions necessary for multiple regression modelling were tested, including multivariate normality, minimal multicollinearity, homoscedasticity, linear relationships between predictor variables and the dependent variable, and normality of dependent variable population. All assumptions were broadly met, with correlations between predictor variables all below 0.5, VIF < 1.4 and tolerance < 1.

In summary, Model 1 indicates that verbs, paternal age and paternal education are all significant predictors of VOCD, with dummy usage of borderline significance, and verbs having the greatest influence. Overall, the model explained 37% of the variance in VOCD.

Model 2: LuCiD Variables with NDW as Outcome Measure

A second stepwise regression was performed incorporating all the LuCiD linguistic and demographic variables, this time with No. of Different Words (NDW) as the dependent variable. No. of Verbs, No. of Rare Words, Gender and Mean Frequency emerged as the predictor variables to include in the model, the results of which are shown below in Table 7.

Table 7

Summary of regression analysis for predicting No. of Different Words

	B	SE(B)	β	t	Sig.	
No. of Verbs	1.67	0.2	0.51	8.38	<0.001	***
No. of Rare Words	2.34	0.54	0.2	4.37	<0.001	***
Gender	-5.46	2.57	-0.07	-2.12	0.038	*
Mean Frequency	-0.002	0.0003	-0.36	-6.63	<0.001	***

No. of Verbs: number of verbs spoken by each participant, No. of Rare Words: number of rare words spoken by each participant, Gender: as determined at birth, Mean Frequency: mean frequency rating for words spoken by each participant.

*. p = 0.05-0.1 * p = 0.01-0.05 ** p=0.001-0.01 *** p=0-0.001*

The predictor variables were shown to explain 92% of variation in NDW (multiple R^2 = 0.92), with the overall model being deemed as significant, $F(4,68) = 204.9$, $p<.001$.

Table 7 shows that No. of Verbs (B = 1.67, $t(68)$ = 8.38, p <.001), No. of Rare Words (B = 2.34, $t(68)$ = 4.37, p <.001), Gender (B = -5.46, $t(68)$ = 2.12, p = .04), and Mean Frequency (B = -0.002, $t(68)$ = 6.63, p < .001) were significant predictors of NDW: as the number of verbs and the number of rare words increased, so did the number of different words (NDW) used by a child. Males were associated with a lower no. of different words, as were children who used on average a higher number of more frequent words.

The standardised coefficients show that a one standard deviation rise in No. of Verbs predicted a 0.51 standard deviation rise in NDW, a one standard deviation rise in No. of Rare Words

predicted a 0.2 standard deviation rise in NDW and a one standard deviation rise in Mean Frequency predicted a -0.36 standard deviation fall in NDW. No. of Verbs was therefore the strongest predictor of NDW in children, with a standardised coefficient of 0.51 indicating a large effect size (Cohen, 1988).

Unstandardised coefficients indicate that with each extra verb spoken, the NDW used increases by 1.67 (range 40-227), and with each extra rare word used, the NDW increases by 2.34. Males are associated with fewer different words, as are those children who use more frequent words: a one unit increase in frequency score (range 22245-57124) is associated with a drop of 0.002 in the number of different words.

The core assumptions necessary for multiple regression modelling were tested, including multivariate normality, minimal multicollinearity, homoscedasticity, linear relationships between predictor variables and the dependent variable, and normality of dependent variable population. All assumptions were met, with correlations between predictor variables all below 0.8, VIF < 3.3 and tolerance < 1.

In summary, Model 2 indicates that verbs, rare words, gender and word frequency are all significant predictors of NDW, with verbs having the greatest influence. Overall, the model explained 92% of the variance in NDW.

Model 3: MCS Variables with BAS as Outcome Measure

For reasons already discussed, a separate analysis was performed using the MCS data to determine the relative influence of demographic variables (see Table 5) on children's vocabulary. Due to a smaller number of predictor variables, this time a multiple regression was performed with BAS t Score as the dependent variable, and No. of Children in Household, Employment Status, Dad Age, Gender, Income, Mum Age, Mum Education and Dad Education as the predictor variables. The results are shown below in Table 8.

The predictor variables were shown to explain 7% of variation in BAS t Scores, with the overall model being deemed as significant $F(8,7643) = 70.43$, $p<.001$).

Table 6 shows that the No. of Children in House (B = -1.38, $t(7643) = 10.1$, $p < .001$), Gender (B = 2.53, $t(7643) = 10.89$, $p < .001$), Income (B = 0.68, $t(7643) = 5.73$, p <.001), Mum Age (B = 0.82, $t(7643) = 3.38$, p =.001), Mum Education (B = 1.03, $t(7643) = 8.35$, p <.001), and Dad Education (B =0.34, $t(7643) = 2.82$, p = .005) were significant predictors of BAS t Scores. Therefore, as income, maternal age, maternal education and paternal education increased, so did scores on the vocabulary subtest of the British Ability Scales (BAS). Males were associated with lower scores, as were higher numbers of children in the home.

The standardised coefficients show that a one standard deviation rise in No. of Children in House predicted a -0.12 standard deviation fall in BAS score, a one standard deviation rise in Income predicted a 0.07 standard deviation rise in BAS score, a one standard deviation rise in Mum Age predicted a 0.05 standard deviation rise in BAS score. No. of children in the house and gender were the strongest predictors of BAS scores in children, followed closely by mum's education, with standardised coefficients of 0.12 and 0.11 respectively, indicating small-medium effect sizes (Cohen, 1988).

Table 8

Summary of regression analysis for predicting BAS t Scores

	B	SE(B)	β	t	Sig.	
No. Children in House	-1.38	0.14	-0.12	-10.1	<0.001	***
Employment Status	-0.27	0.23	-0.01	-1.17	0.24	
Dad Age	0.14	0.23	0.01	0.6	0.55	
Gender	2.53	0.23	0.12	10.89	<0.001	***
Income	0.68	0.12	0.07	5.73	<0.001	***
Mum Age	0.82	0.24	0.05	3.38	0.001	***
Mum Education	1.03	0.12	0.11	8.35	<0.001	***
Dad Education	0.34	0.12	0.04	2.82	0.005	**

No. of Chn in House: number of children in household including participant and siblings, when participant aged 36 months, Employment Status: combined employment status of parents/parent + partner when participant aged 36 months, Dad Age: dad's age in years at participant's birth, Gender: as determined at birth, Income: income measured when participant aged 36 months, Mum Age: mum's age in years at participant's birth, Mum Education: mum's highest level of education when participant aged 36 months, Dad Education: dad's highest level of education when participant aged 36 months.

*. p = 0.05-0.1 * p = 0.01-0.05 ** p=0.001-0.01 *** p=0-0.001*

Unstandardised coefficients indicate that with each extra child in the household, BAS score decreased by 1.38 (range 20-40), whereas an increase into the next income bracket had an associated 0.68 increase in BAS scores, and an increase into the next age group for mothers was associated with a 0.82 increase in BAS scores. A one unit increase in qualification level was associated with a 1.03 increase in BAS scores for mothers, and 0.34 for fathers.

The core assumptions necessary for multiple regression modelling were tested, including multivariate normality, minimal multicollinearity, homoscedasticity, linear relationships between predictor variables and the dependent variable, and normality of dependent variable population. All assumptions were met, with correlations between predictor variables all below 0.6, VIF < 1.6 and tolerance < 1.

In summary, Model 3 indicates that no. of children in a household, gender, income, maternal age, maternal education and paternal education are all significant predictors of BAS vocabulary scores, with no. of children and gender having the greatest influence. Overall, the model explained 7% of the variance in BAS scores.

Secondary Analysis

Due to the results of Models 1 and 2 alluding to the relative importance of verbs (verbs emerged as the only common significant predictor across both models for different measures of lexical diversity, and in both models, verbs were the most influential predictor) further regression analysis was carried out. This was done using a subset of each transcript containing only utterances that included a specific set of ten common verbs, with the aim to ascertain whether or not those children who used a greater number of different verbs (= more 'verbally diverse') tended to have greater lexical

diversity in the utterances which contained a particular set of verbs, as opposed to a less verbally diverse child using the same set of verbs – in other words comparing how different children used the same verb within an utterance. The results of such an analysis would shed light on whether children who used more verbs tended to employ additional word types around those verbs, hence increasing their overall lexical diversity, and suggesting that it was the verbs themselves that were driving the extra acquisition of vocabulary.

Model 4: Number of Verbs with VOCD as Outcome Measure

A linear regression was performed, with VOCD of subset as the dependent variable, and No. of Verbs as the predictor variable. The results indicated No. of Verbs to be a significant predictor (p = .001) of VOCD (subset), explaining 15% of the variance in the outcome variable.

As No. of Verbs (measuring each child's overall verbal diversity for the whole transcript) increased, so did the VOCD of the subset, indicating that the more verbally diverse children (overall) employed more non-verb word types around the same set of ten common verbs than when compared to the less verbally diverse children (overall). It is possible that the predictive link between No. of Verbs and lexical diversity evidenced by Models 1 and 2 could have simply been due to the logic that more diversity in verb usage feeds into more diversity overall, however the results of this secondary analysis suggest that more diversity in verb usage is instead driving the acquisition of additional non-verb vocabulary to be used alongside those verbs.

The core assumptions necessary for linear regression modelling were tested, including multivariate normality, homoscedasticity, linear relationship between predictor variable and dependent variable, and normality of dependent variable population. All assumptions were broadly met.

Model 5: Number of Verbs with NDW as Outcome Measure

In order to mirror the outcome variables used in the main analysis, a final linear regression was performed, this time with NDW (subset) as the dependent variable, and No. of Verbs as the predictor variable. The results indicated No. of Verbs to be a significant predictor (p = .005) of NDW (subset), explaining 6% of the variance in the outcome variable. NDW (subset) used an adjusted score which accounted for the varied token size of each subset transcript (no longer standardised by the ten-minute time window as the full transcript was), and was in the form of a type-token ratio.

As No. of Verbs (measuring each child's overall verbal diversity for the whole transcript) increased, the adjusted NDW score in the subset decreased, indicating more verbally diverse children used less word types. This result, contrary to the results of Model 4, is unexpected and will be fully explored in the following chapter.

The core assumptions necessary for linear regression modelling were tested, including multivariate normality, homoscedasticity, linear relationship between predictor variable and dependent variable, and normality of dependent variable population. All assumptions were broadly met.

Chapter 6: Discussion (Phase One)

Aims and Hypotheses

The aim of Phase One of this study was primarily to investigate the expressive vocabulary differences evident in pre-school children in a natural setting (using the LuCiD dataset) with a view to ascertaining whether vocabulary (measured in terms of lexical diversity) could be predicted by particular linguistic and demographic measures. Research Objective 1 was therefore to explore the linguistic characteristics of the vocabulary gap (the difference between largest and smallest), and Research Objective 2 to investigate the most salient demographic risk factors associated with smaller vocabularies at pre-school age. Concerning linguistic characteristics, whilst the nature of the vocabulary gap (in terms of its linguistic character) is unclear, using age of acquisition (AofA) research, it was hypothesised that word frequency, lexical category, word length, concreteness, phonological neighbourhood density and phonotactic probability would all be significant predictors of vocabulary (measured in terms of lexical diversity, in two different ways). Concerning demographics, it was hypothesised that income, maternal and paternal age and education level, size of household, gender, dummy usage and employment status (the latter only for the MCS dataset) would all be significant predictors of vocabulary (measured in terms of lexical diversity and BAS scores) at age 3. The expected influence of paternal caregiving responsibilities was unclear, and the variable was included for interest due to the shifting balance of research towards paternal influences on children's outcomes to reflect the changing demographics of work and family life in the 21st century.

Key Findings

The results of the regression analyses supported some of the hypothesised predictions and not others. Lexical diversity was measured in two ways - vocabulary diversity (VOCD) and number of different words (NDW). When VOCD was the measure of lexical diversity, the number of verbs used by the child, paternal age and paternal education were all positive significant predictors, with dummy usage of marginal significance. When NDW was the measure of lexical diversity, the number of verbs, the number of rare words, the word's overall frequency and gender were significant predictors, with the former two variables predicting NDW in a positive direction, and word frequency having a negative influence (i.e., children who used a higher proportion of highly frequent words tended to have lower NDW scores). Boys also tended to have lower NDW scores.

Research Objective 1

Analysis of the LuCiD data indicates that the linguistic character of the vocabulary gap can be reliably defined in terms of the number of verbs used by the child, which was consistently found to be a significant predictor of a child's overall vocabulary diversity. The results of a secondary analysis, designed to ascertain the nature of the role played by verbs by comparing how different children used the same verb within an utterance, showed that the number of verbs a child used overall was a significant predictor of VOCD in a smaller subset transcript where only utterances that contained specific common verbs were included – indicating that more 'verbally diverse' children employ more additional non-verb word types around their verbs, suggesting that it is the verbs themselves that are facilitating higher levels of lexical diversity. This result was different however when NDW was used as the measure for vocabulary diversity, suggesting more verbally diverse children use less word types

to accompany the verbs they used, hence both results should be interpreted with caution, and will be explored more fully later on.

The number of rare and less frequent words used also appears to be a significant predictor of a child's overall lexical diversity, although discrepancies between the results obtained for the two different measures of lexical diversity (VOCD and NDW) necessitate caution. Put more simply then, 3-year-old children who have smaller vocabularies when compared to their more expressive peers appear to be lacking in verbs, and possibly rare and less frequent words, and therefore may benefit from interventions designed specifically to teach these word types.

Research Objective 2

Analysis of the LuCiD data indicates that paternal age, paternal education and gender are the most salient demographic risk factors associated with smaller vocabularies at age 3, with a possible role for dummy usage (in that higher usage was borderline predictive of lower vocabulary scores). Again, caution should be employed when interpreting these results due to the discrepancy between the results obtained for the two different measures of lexical diversity (VOCD and NDW). This will be explored in more detail below.

Similar analyses conducted on the MCS dataset for triangulation purposes broadly supported the stated hypotheses for Research Objective 2, with the number of children in the household, gender, income, maternal age, maternal education and paternal education emerging as significant predictors of vocabulary (in this case measured by BAS scores), and therefore salient demographic risk factors associated with smaller vocabularies at age 3. The MCS analysis indicated a broader range of demographic risk factors for smaller vocabularies than the LuCiD dataset, and thus the apparent differences between the two datasets will be considered below.

Summary of Key Findings

To summarise the overall results, vocabulary size (or lexical diversity) is consistently predicted by the number of verbs a child uses, and also by the number of rare and less frequent words they use (but only for one of the measures of lexical diversity). The demographic risk factors for a smaller vocabulary at age 3 have been shown to be gender, paternal and maternal age, income, the number of children in a household, maternal and paternal education, and possibly dummy usage. Discrepancies between the results - in terms of both the measures of vocabulary used, and the different datasets - require further consideration.

The Linguistic Character of the Vocabulary Gap (Research Objective 1)

Research Objective 1 aimed to explore the linguistic characteristics of the vocabulary gap, in terms of the types of word which may reside within the gap, and thus inform the content of interventions which aim to close the gap, and to support children's overall language development. This objective was addressed by analysing a corpus of children's talk against a selection of linguistic criteria shown by existing evidence to be important predictors of ease of acquisition.

Verbs

This study has shown that the number of verbs a child expresses in a naturalistic setting is predictive of their overall lexical diversity in the same setting (regardless of the outcome measure used: VOCD or NDW). This finding indicates that children who use more verb types also tend to have larger/more-diverse vocabularies as hypothesised, suggesting that verbs are likely to form part of the vocabulary gap. Additional analyses tentatively indicates that the extra words used by children with larger more diverse vocabularies is made up not just of those additional verbs, but also of other word types. A focus on the teaching of verbs, therefore, may facilitate overall vocabulary growth for those children with smaller vocabularies

Verbs are Acquired Later. These results appear to be supported by the current literature, if somewhat indirectly. Age of acquisition research has shown that verbs are generally considered more difficult to learn than other word types, and tend to be learned later. This corresponds with the idea that they are associated with a later developmental stage and a larger more diverse vocabulary. In general, and across languages, nouns tend to be acquired before other lexical categories (Goodman et al., 2008), possibly because they have easier, everyday referents (Baldwin, 1995), and such referents tend to be more concrete (Hansen, 2017). Verbs appear later and are initially harder to learn because acquiring them involves working out the complex relationship between form, function, culturally relevant word meaning and linguistically relevant relational meaning (Braunwald, 1995). Their form can change depending on the context, they are often more abstract (Gentner, 2006), and do not always have a clear beginning and end point, making them more difficult to dissect from the overall scene (Abbot-Smith et al., 2017). The fact that verbs appear later does not on its own explain why they predict lexical diversity, but it does show a general synchronicity between verb usage and larger vocabularies in young children.

Verbs go Alongside the Ability to Produce Multi-word Utterances. As well as being associated with a later developmental stage, verb proficiency has also been positively linked with children's grammar skills. Hadley et al. (2016) found that the number of verbs produced by 2-year-olds, both via parent report and in samples of spontaneous speech, predicted later grammar skills as measured by the Index of Productive Syntax, with the latter being the strongest predictor. A limited verb lexicon, they argued, should serve as a red flag to parents and professionals for developmental language disorders later on, but could also serve as a red flag for more general language delays, and a marker for early intervention. There appears to be a general interdependence between lexical and syntactic development (Tomasello, 2005; Braginsky et al., 2015), and this inherent link between grammar and vocabulary suggests that if verbs have predicted grammar skills in previous research (e.g., Hadley et al., 2016), that they have also predicted lexical diversity in the current research is not unsurprising.

Once acquired, verbs provide the vital component for multi-word utterances (Horvath et al., 2022), and due to their relational function linking the different parts of an utterance, often appear shortly before or alongside the ability to produce simple subject-verb-object sentences (Braunwald, 1995). Whilst verbs can be used in isolation to communicate meaning, once a child begins to utilise them in sentences, a wider selection of supporting vocabulary is required in order to communicate more complex meanings. Such vocabulary may form part of a child's existing lexicon, or it may be words as yet unknown: it thus seems logical to speculate, specifically in the latter case, that the acquisition of

verbs and the ability to produce multi-word utterances has been the driver of further vocabulary acquisition.

Verbs Appear to Drive Additional Vocabulary Acquisition. The importance of verbs for vocabulary development appears to go beyond their association with superior grammar and their functional role within a sentence. By carrying out additional analysis designed to ascertain the nature of the role played by verbs, the findings in the current study have suggested that it could be the verbs themselves which appear to be driving the additional non-verb vocabulary acquisition, a position that is supported by the preceding paragraphs. This additional analysis compared how different children used the same verb within an utterance and showed that the number of verbs a child used overall was a significant predictor of VOCD in a smaller subset transcript where only utterances that contained specific common verbs were included – in other words, those children that tended to use more verbs in general would also tend to employ more supporting non-verb word types around their verbs (thus resulting in a greater overall lexical diversity).

What Comes First – Verbs or Grammar? Results of the current study have shown that the number of verbs a child produces is predictive of their overall lexicon, a result that is confirmed or alluded to by other studies. Further analysis was consistent with the notion that the verbs themselves appear to be driving the additional non-verb vocabulary acquisition (possibly due to the requirement for more supporting non-verb word types in utterances containing verbs), and so the predictive relationship between verbs and vocabulary diversity is over and above the logical conclusion that greater verb diversity necessarily contributes to overall lexical diversity. Yet despite a clear association between the two measures, it remains unclear whether the verbs themselves are a prerequisite and facilitate the acquisition of diverse vocabulary as suggested, or if it is the other way around, and the verbs are necessitated by an influx of new vocabulary in a drive to produce more complex multi-word utterances - or indeed if these things happen to occur synchronously.

The notion that verbs come first, before both the ability to produce sentence-verb-object (SVO) sentences, and before the influx of vocabulary needed to support those sentences, appears to be supported by Braunwald (1995). Here, a detailed study of the language acquisition of two sisters revealed that whilst both children acquired verbs in somewhat different ways (e.g., the firstborn's acquisition of verbs coincided with a general vocabulary spurt and multi-word utterances; the second born acquired verbs in a more idiosyncratic way, and did not use them within SVO sentences for some time) – both girls nonetheless acquired verbs as a prerequisite to producing SVO sentences, even if the time lapse involved differed.

An alternative perspective is that offered by Gentner and Boroditsky (2001), who argue that children need some knowledge of grammar to learn verb meanings, as their meanings are shaped by the language system and cannot exist entirely independently from it. Whilst this position does not assert one way or another whether verbs are a useful pre-requisite to wider vocabulary learning or vice verca, it does suggest the learning of verbs, syntax and other word types are not easily extricated from one another, and do not occur in a neat linear fashion.

One theory that presents a more flexible approach and allows for individual differences between learners is the hybrid Emergentist Coalition Model (Hollich et al., 2000), which asserts that multiple cues are available to children when learning words (including verbs), including perceptual salience,

social and linguistic cues. Actions that are most salient to the child (such as ones they perform themselves) are likely to be learned first, followed by social cues guiding the learner to the intentions of the speaker, and finally linguistic cues whereby the familiar order of words is used to ascertain meaning (syntactic bootstrapping). Whilst children tend to use these cues developmentally in the order shown (suggesting that grammar knowledge - indicated via the use of syntactic bootstrapping - is learned later than initially salient verbs), the shifting combinations of each cue's usage and how they interact together remains unclear (Golinkoff & Hirsh-Pasek, 2008), and it is likely that perceptual salience continues to play a dominant role even as children get older (Abbot-Smith et al., 2017). Whilst the ECM is concerned with how children acquire receptive vocabulary, it can still be used as a framework for understanding vocabulary acquisition overall (e.g., Maekawa & Storkel, 2006), and the mechanisms through which receptive vocabulary knowledge may influence or be translated into expressive vocabulary were discussed in Chapter 1.

In summary then, although the exact timelines will depend on the child and circumstances in question, it is thus likely that children learn some initial verbs alongside other words, then gradual recognition of grammar rules allows them to learn more verbs, which drives the need for more vocabulary and the eventual production of multi-word utterances. This supports the notion, as has been argued here, that verbs drive the acquisition of new vocabulary, and yet it also acknowledges that the process of verb acquisition, grammar learning and vocabulary acquisition in general is an intertwined and reciprocal process, with the implication that future interventions that focus on the teaching of verbs do not present them in isolation.

Are Some Types of Verb More Important than Others? The type of verb might be an important consideration in terms of the strength of its influence on vocabulary acquisition. When investigating the differences between typically developing children and late talkers, the former group having, by definition, larger expressive vocabularies, Horvath et al. (2022) found that typically developing children tended to produce more of a particular type of verb termed a 'manner' verb, as well as more verbs overall. Manner verbs refer to modes of action, for example *scribble*, *sweep* and *run*, whereas 'result' verbs refer to the end state of an event, for example *clean*, *fill* and *arrive.* He at al. (2020) also found that typically developing toddlers were more likely to produce manner verbs than result verbs, and suggested that this might be because such verbs were better suited to support the building of their semantic network, as they allowed or promoted the acquisition of words that connected to vocabulary items already known. Result verbs may also be less visually salient, and therefore young children may be more biased towards mistakenly mapping them onto unknown objects (Abbot-Smith et al., 2017). The current study analysed verbs as a whole class, rather than by type, so cannot corroborate these findings or otherwise. Nonetheless the implications of these results in the context of formulating intervention content and design suggests another important focus for future research.

Differences in the Way Verbs are Defined and Counted. Determining what should count as a verb, which can vary between meaning, function and form, is not a straight-forward and agreed process. Indeed, the differing rules applied by different researchers and in different contexts regarding how verbs should be categorised has been labelled a 'fundamental problem' (Braunwald, 1995, p.81), and necessarily has the potential to dilute the utility of making comparisons between the results of different studies, as well as real-world applicability. For example in Hadley et al. (2016), only open class lexical verbs were included – in that they must carry semantic meaning and be capable of

inflection for tense, with inflectional variants treated as a single type. Irregular past tense and third person singular verbs, gerunds and closed class auxiliary verbs were excluded, although the reason for this is unclear. Braunwald (1995), on the other hand, distinguished between those who defined verbs relying on their meaning and function (and therefore disregarding their form) and those who focused on form, regardless of meaning/function.

With due consideration given to both the pragmatic research paradigm and the practical aims of the current study (to discover the characteristics of the vocabulary gap, find out who is most at risk, and formulate an effective intervention to support those children), and acknowledgement of the complexities involved in categorising verbs, a decision was made to use a broad classification of 'verb' here. BBC Bitesize for key stage 1, an online teaching and learning resource commonly used by children, parents and teachers, describes a verb as 'a word used to describe an action, state or occurrence', with examples of *jumping*, *like* and *became* given. All of these examples can be conceptualised as actions or 'doing words': an understanding of 'verb' that tends to be echoed across key stakeholder groups - parents of young children and adults working in early years settings. If this broad categorisation of doing/action is how verbs are understood amongst these important real-world stakeholders in children's language development, then it is imperative that the way the current research defines and categorises verb is reflective of this. Therefore, for the current study, words such as *playing* (gerund), and different forms of the same word (e.g., *play* and *played*) were all included as valid and separate verbs. This decision will be revisited as part of a wider discussion below with respect to the different measures of lexical diversity: NDW and VOCD.

To operationalise the broad classification of verb used for this study, the MOR programme (built into the CLAN software and intended for morphosyntactic analysis of speech samples in the CHAT format) provided parts of speech tagging by assigning each word spoken into one of 39 parts of speech categories, with additional sub-categories. 'Verbs' were considered to be any word that had been categorised as 'verb', 'auxiliary verb', 'modal verb', 'copula verb' or 'participle verb'. Different forms of a verb (e.g., *fish* and *fishing*, *do* and *does*) were counted as different verbs, and words that could be classed as nouns or verbs (e.g., *drink*) were examined in context on a case-by-case basis to ascertain intended meaning.

The inevitable consequence of using a wider classification of verb in the current study is that other studies that have used a different definition of verb will necessarily become less comparable. On the other hand, and most importantly for the pragmatic aims of this research, a broad classification reinforces the applicability of these results to their associated real-world context, in that the way verbs have been defined and categorised is reflective of how key stakeholders in children's language development also define verbs. In lieu of existing data to support this assertion, this decision was based on existing knowledge of literacy teaching and learning resources.

Verbs can also be counted and accounted for statistically in different ways (indeed this is true of any categorical indicator applied to a word). In the current study, raw counts of verbs were used, similar to NDW and number of rare words, with the view that the speech sample had been standardised by time (recording duration being equal across participants), so that all children had an equal amount of time to express the verbs that they knew. It would have been possible to account for verbs as a proportion of the total number of words spoken by the child, however this was discounted

to avoid issues of using proportion data that adds up to one, and therefore is not linearly independent. The same approach was used by Rowe (2012) with respect to rare words, and Hadley et al. (2016).

Summary: Verbs and the Vocabulary Gap. This study has shown that children's use of verbs (in terms of the number of different types of verbs they use) is predictive of their overall expressive vocabulary in a naturalistic setting (regardless of the outcome measure used). It has also shown that lexically diverse children tend to employ more non-verb word types around their verbs, suggesting that verb usage may be facilitating additional vocabulary acquisition, and that the predictive relationship between verbs and vocabulary diversity is over and above the logical conclusion that greater verb diversity necessarily contributes to overall lexical diversity. One explanation for these findings is that verbs do not always communicate meaning effectively on their own, and tend to reside within multi-word utterances, thus requiring additional verb types around them. Verbs also tend to be more difficult to acquire and thus acquired later – therefore it makes sense that greater verb diversity corresponds with overall language development trajectories, and those experiencing language delay may need particular support with verb learning.

Verb usage is also inextricably wound up with grammatical knowledge, possibly explained by a verb's vital relational role within multi-word utterances. The direction of the relationship between both verb use and additional vocabulary, and indeed verb usage and grammatical knowledge, was also considered, and it was argued that whilst, in support of the results obtained here, verbs appear to facilitate the acquisition of new vocabulary, it also likely that the process of verb acquisition, grammar learning and vocabulary acquisition in general are an intertwined and reciprocal process. Whether or not the type of verb might be relevant was also briefly discussed, and consideration was given to the way verbs are defined, counted and statistically analysed, and any implications this might have. Bearing in mind the pragmatic research paradigm and the practical aims of the current study (to discover the characteristics of the vocabulary gap, find out who is most at risk, and formulate an effective intervention to support those children), a broad classification of 'verb' was justified. Ultimately, a focus on the teaching of verbs within intervention contexts may facilitate overall vocabulary growth for those children with smaller vocabularies, particularly if those interventions pay heed to the intertwined and reciprocal relations between verbs, overall vocabulary and grammatical knowledge.

Word Frequency and Rare Words

This study has shown that the number of rare words and the number of less frequent words a child expresses in a naturalistic setting is predictive of their overall lexical diversity in the same setting, indicating that children who use more less frequent words, and a higher proportion of rare words, also tend to have larger more diverse vocabularies as hypothesised. Less frequent and rare words are therefore likely to form part of the vocabulary gap. This is only the case, however, when NDW is the outcome measure. 'Rare' status was conferred upon any word that was not considered known (as judged by teachers) by any more than 20% of US 4th graders as per the Dale-Chall word list, and word frequency referred to how commonly a word appears in the language, in this case calculated via the spoken BNC.

Less Frequent/Rare Words are Acquired later and Via Multi-directional Influence . As with verbs, less frequent and rare words tend to be acquired later than more frequent everyday words, and are

therefore associated with a later developmental stage. In line with constructivist frameworks of language acquisition, the first words children acquire tend to be the ones that occur more frequently in the language (Swingley & Humphrey, 2018), as these are the words that they hear the most from their caregivers, plausibly because they are of particular salience to the child (e.g., *mum*, *dad*, *dog* etc). If a child is never exposed to a particular word, then there is no way of them learning it, as word labels and meanings cannot exist innately (Ambridge & Lieven, 2011), and depend upon the language being spoken. The fact that less frequent and rare words tend to appear later does not on its own explain why they predict lexical diversity, but it does show a general synchronicity between rare word usage and larger vocabularies in young children.

Constructivist paradigms of word-learning maintain that children construct vocabulary knowledge from their environment, a large component of which is the verbal input they receive from parents. Evidence suggests that parents' use of rare words with their 30-month-olds positively predicted children's vocabularies one year later (Rowe, 2012), and that there is a multi-directional influence between CDS and infant talk (Huttenlocher et al., 2010), with each predicting the other and resulting in a spiral of accumulating advantages for children who receive quality input. The latter study was concerned with diversity of vocabulary rather than the quantity of rare/infrequent words used, however the same logic can be applied: children who are exposed to rare words subsequently begin to express those words, and their caregivers respond to this signal of readiness by using additional rare words, resulting in even further vocabulary being learned, and so on. This offers some explanation as to why rare or less frequent words may predict overall vocabulary size. A further explanation is that, due to the nature of language in books versus everyday speech, rare words tend to be more prevalent in shared reading contexts, and so a child who is exposed to and subsequently acquires a greater proportion of rare words than their peers could also be experiencing more shared reading interactions with their caregiver, something which is independently and positively related to their language development (Pillinger & Wood, 2014).

Frequency Plays a Complex Role in Word Acquisition. The current results partially support the notion that words which are less frequent in the language are harder to learn and acquired later (in that word frequency significantly predicted vocabulary diversity, but only when measured via the NDW metric). This conflicting result is somewhat reflected by wider research in which the importance of word frequency in language acquisition is supported, but not unequivocally without qualification (Stokes, 2010). Goodman et al. (2008) found that frequency was only associated with earlier acquisition within specific lexical categories – and not when the whole vocabulary was considered. For example, closed-class words (such as pronouns, determiners and conjunctions) were the most frequently occurring words individually, yet often acquired later. At the other end of the scale, nouns occurred least frequently individually, but were learned the earliest. When nouns were considered on their own however, the frequency with which a noun was heard was important for how early it was acquired. Thus, whilst there is a general consensus that frequency is relevant with respect to which words children learn first (Rowe, 2012; Weizman & Snow, 2001), there are many other factors at play that may qualify or interact with frequency, including PND (Stokes, 2010), the quality and circumstances of the exchange, for example joint attention (Tomasello & Haberl, 2003) contingent talk (McGillion et al., 2017) whether the talk is embedded in useful and instructive interactions (Weisman & Snow, 2001) the appearance of the word in isolation (Brent & Siskind, 2001) or in utterance final position (Naigles & Hoff-Ginsberg, 1998), lexical category and developmental stage (Goodman et al., 2008).

PND (Phonological Neighbourhood Density), referring to the number of words that can be formed by addition, substitution or elimination of a single phoneme in the target word, appears to be a word characteristic that is particularly interwoven with word frequency. Stokes (2010) suggested that PND and word frequency were so highly correlated that limited conclusions could be drawn regarding frequency as a predictor of vocabulary, as words that were highly frequent also tended to have a high number of phonological neighbours. This was likely to be the underlying reason for the early acquisition of such words, as well as a possible explanation for why less frequent, rare words tend to be acquired later on and associated with larger vocabularies: exposure to many common words with many phonological neighbours allows a large store of sub-lexical chunks to be accumulated, which then facilitates less frequent words, with less phonological neighbours, to be acquired. Once this process has begun, a 'virtuous circle' (Jones & Rowland, 2017, p.19) commences, whereby a wide variety of sub-lexical knowledge can be called upon to accelerate the learning of new words, including increasingly complex words.

How Measures of Word Frequency and Rareness Reflect the Input Children Hear. As noted in Chapter 3, it is important to reconsider the measures used to operationalise rare word status and word frequency, and acknowledge the potential limitations in their ability to reflect the words children themselves are likely to hear. In other words, how likely it is that a specific word rated as both rare and relatively infrequent (by reference to the Dale-Chall and spoken BNC subset respectively) would also be considered so within the child's own context, and specifically regarding the words which they are exposed to via CDS? If these measures did not correspond well with actual verbal input, what impact might this have had on the results obtained?

The Dale-Chall wordlist was used to operationalise word rareness, as was the case in Rowe (2012) and Weizman and Snow (2001). As noted in Chapter 4, the list contains 3000 common words which are known (as judged by teachers) in reading by most 9 to 10-year-olds in the USA. Whilst the measure is grounded in the reading knowledge of older children (thus some words may be less likely to be considered rare than for a younger child), it is still a child-focused measure, which one might presume would more accurately reflect the language used in CDS with younger children than an adult-focused measure of frequency, although this would depend on individual factors such as the type of books being read to the child. On the other hand, again as detailed in Chapter 4, this method of determining rareness did require manual checking and occasional removal of a word's rare status on a case-by-case basis (e.g., *mummy* and *teddy* were removed from the rare list). The necessity of this process does suggest some slight limitations with the Dale-Chall approach, in that the list alone was not completely reflective of rare words within the context of infant speech. This could have been due to cultural differences between US and UK language norms (e.g., *mum* versus *mom*), as well as the absence from the list of particular words (presumably due to their reduced usage amongst older children) which are common in infant interactions (e.g., some animal noises, which are a distinct category within the CDI), therefore erroneously categorising them as rare. By making the described amendments however, an attempt was made to make the measure more reflective of what could sensibly be considered rare within UK infant speech contexts.

Word frequency was measured here with reference to a subset of data from the spoken BNC, as was used in Jones et al. (2021). This included speech from unscripted informal conversations and a range of other, more formal, contexts, including classrooms, business meetings and sports

commentaries, and was balanced across different age-groups and backgrounds. Despite the balance, the BNC measure was nonetheless a more adult-focused estimate of word frequency than one taken, for example, from transcripts of CDS. This means that a word with a high frequency rating from the BNC, might not be a highly frequent word in the context of infants' interactions with their caregivers. Indeed, Goodman et al. (2008) incorporated different measures of word frequency, including both child-focused measures (using CDS from CHILDES) and adult-focused/older children-focused measures (using Kucera–Francis and Thorndike–Lorge norms), and found that the effects of frequency on age of acquisition for specific words were much stronger when frequency was measured by the CHILDES CDS measure than by the older-child/adult-focused measures. Jones et al. (2023) also found that the effects of frequency on children's productive vocabulary learning were stronger when frequency was calculated from CDS corpora as opposed to more adult-focused measures. Here, the differing effects of CDS versus adult-directed or overheard speech on children's productive vocabulary were explained via different characteristics of these two types of speech, with CDS containing a higher proportion of words which were short, frequent, had high neighbourhood densities and low phonotactic probabilities. CDS also contains characteristics such as exaggerated prosody and contingency which are thought to assist a child's vocabulary learning (Golinkoff et al., 2015). On the other hand, a difference in strength of effect between measures does not equate to no effect at all when using more adult-focused measures of frequency, and both Jones et al. (2023) and Goodman et al. (2008) found a less pronounced effect when adult-focused frequency measures were used. Indeed, Stokes (2010) found a strong positive correlation between the CHILDES measure of word frequency and the adult-focused De Cara and Goswami frequencies (r=.90), and within their own analysis similar results were found for both measures in terms of the influence on children's vocabulary learning.

The current study included two similar measures of how commonly a word appears in the language – a measure of rareness and frequency - both of which were established to be important predictors of word learning within the existing literature (Rowe, 2012; Swingley & Humphrey, 2018). Whilst it seems sensible to presume that rare words will also be infrequent, this will not necessarily be the case due to the differing ways the measures are operationalised, as described above. Whilst the limitations of these measures have been acknowledged above in terms of their ability to truly reflect the verbal input a child receives, by including both measures, this allowed for different elements of a word's commonality to be captured, particularly as one measure involved a more child-focused measure, and the other was more adult-focused in nature.

Summary: Frequency and Rare Words. This study has shown that the number of rare words and the number of less-frequent words a child expresses in a naturalistic setting are predictive of their overall lexical diversity, suggesting that less frequent and rare words may form part of the vocabulary gap. This was only the case when NDW was used as the outcome measure however, and this conflict within the results is reflected in the broader existing research which qualifies the importance of word frequency, for example within lexical category, PND or quality of the communication exchange.

In a similar vein to verb usage, one explanation for less-frequent/rare words being less prevalent within smaller vocabularies is that they tend to be acquired at a later developmental stage, as children tend to learn the words which they hear the most first (and these words are also likely to be the most salient to them). As children come to know and use a greater proportion of rare/less-frequent words, this could provide a signal to their caregivers demonstrating their readiness for receiving even further

rare word input, thus benefitting the child further. This is similar to the reciprocal relationship found in Huttenlocher et al. (2010) whereby diversity of earlier caregiver speech predicted later child speech, and earlier child speech also predicted later caregiver speech. Rare word usage could also indicate a high degee of shared reading, which is independently and positively related to children's language development.

Consideration was also given to the limitations of the measures used in terms of their ability to reflect the verbal input children receive, with some studies suggesting more pronounced effects will be found when employing more child-focused measures of frequency. Whilst acknowledging these limitations, including two measures in the current study was intended to allow for slightly different elements of a word's commonality to be captured, particularly as one measure involved a more child-focused measure, and the other was more adult-focused in nature.

Non-significant Variables

Results for the current study indicate that the mean number of phonemes, number of syllables, concreteness rating, PND and PP of words produced by children in a naturalistic setting does not predict their overall lexical diversity in the same setting, or there is insufficient evidence to confirm such a relationship. The number of adjectives produced was similarly non-significant. This suggests that none of these variables influence which word types may be present within the vocabulary gap.

Number of Phonemes and Number of Syllables. With number of phonemes and number of syllables intended as proxies for word length and pronunciation difficulty, the non-significant results suggest that these variables are not predictive of vocabulary size, and therefore are of limited relevance to the vocabulary gap. The inclusion of these variables, and the associated initial hypothesis (that longer and more difficult to pronounce words would predict greater vocabulary diversity) was based on studies showing that the number of phonemes (as a marker of pronunciation) and the number of syllables (measuring length) were predictive of early acquisition (Frank et al., 2021; Jones et al., 2021). These results are difficult to reconcile with those of the current study, but it is important to point out that Frank et al. (2021) used CDI data as opposed to naturalistic data, and in Jones et al. (2021) the data was retrieved from younger children and over a longer period of one year, thus involving considerably more hours of talk (408 hours for 12 children as opposed to just over 24 hours for 73 children in the current study) – indicating possible sample size limitations.

On the other hand, the overall position is rendered unclear by the fact that not all studies have found length (in terms of characters or syllables) to be important for ease of acquisition (Braginsky et al., 2016; Swingley & Humphrey, 2018). The position is further confused by the fact that both word length and pronunciation difficulty are operationalised via the same measures: phonemes, characters and syllables, all of which are highly correlated (Lewis & Frank, 2016). Indeed, ease of pronunciation is particularly difficult to draw conclusions about due to the difficulty involved in its measurement. Despite a clear association between ease of pronunciation and word length (with longer words being more difficult to pronounce), obviously this is not always the case, and may depend more on the specific combination of phonemes, rather than how many there are. This sheds doubt on whether it is appropriate to use the same metrics to measure both length and ease of pronunciation, and a more specific measure, perhaps involving the presence of particular consonant clusters, is required (although such a measure has yet to be developed for use at scale).

In summary, there was insufficient evidence to show a significant relationship between either phonemes or syllables and vocabulary diversity. Again, existing research is somewhat contradictory, suggesting word length is likely to be an important factor in how children acquire language, but will probably interplay with other important factors such as PND and frequency. It seems plausible that ease of pronunciation should also be an important factor (with words that are more complicated to pronounce being more likely to reside within the vocabulary gap), however better measurement tools and further research are required before such a relationship can be confirmed.

Concreteness. There was insufficient evidence to link a word's concreteness rating with overall lexical diversity, indicating that more abstract words are not necessarily associated with larger, more diverse vocabularies, and therefore not necessarily part of the vocabulary gap as hypothesised. This hypothesis was based on support for the idea that a word's concreteness could affect its ease of acquisition with respect to production (Jones & Brandt, 2019; Braginsky et al., 2016), and yet there are some qualifications to this idea that may help to reconcile the current results with those of existing research. For example, in Braginsky et al. (2016), concreteness demonstrated a U-shaped trajectory in its importance throughout development, with the earliest words being relatively abstract, and concreteness becoming more important at around 14 months old, and then becoming less so as time went on, although data was also only available up to 30 months old – as well as being concerned with children's comprehension as opposed to production. In Swingley and Humphrey (2018) concreteness was again found to be predictive, but at an earlier age of 15 months, with a different measure of concreteness and a different vocabulary outcome measure, and only for comprehension not production. It was suggested that this may be because concreteness predicted the ease of learning the meaning of a word for receptive vocabulary, but for producing a word phonological factors become more important (although this was not found to be the case in their study, with PP and number of syllables not found to be significant).

In summary then, concreteness was not shown to be an important factor in predicting vocabulary diversity for this particular sample, and therefore its relationship with language learning and the vocabulary gap is unclear. The non-significant finding in this study might be explained by the age of the current sample, in that the influence of concreteness is dependent on developmental stage as per the U-shaped relationship suggested by Braginsky et al. (2016) with respect to receptive vocabulary, although design and sample differences in general make direct comparisons with existing studies somewhat flawed.

Phonological Neighbourhood Density (PND) and Phonotactic Probability (PP). It was hypothesised that words with sparse phonological neighbourhoods and words with high PP would be associated with larger, more diverse vocabularies, and would be likely to form part of the vocabulary gap. This was not found to be the case. Despite the initial hypotheses, the balance of evidence on both PND and PP is ambiguous, in that their importance appears to be qualified by other factors such as a child's age and word length. Jones and Brandt (2019) found PND to be significant for word production in children aged 12-25 months old, but, crucially, more important for 12-month-olds than for 25-month-olds, indicating a downward trend in influence that may help to explain the results of the current study in relation to 36-month-olds. The data used in the former study was also CDI data as opposed to natural language samples, which may provide further explanation for the contradictory results.

It has also been argued that word length may confound the effects of PP. In Jones et al. (2021), whilst words with a high PND were learned first, the effects of PP were confounded by word length: if length was not controlled then words with low PP tended to be learnt first simply because they were shorter. Indeed, it was argued that the 'varied grain size of phonological knowledge at the point of learning' (Jones et al., 2021, p.11) was the important metric over and above PND and PP. In summary, the relationship between PP and PND and the vocabulary gap is unclear, with their influence on vocabulary acquisition likely to be affected by other word characteristics, and possibly developmental stage. Future research should look to unpick these relative influences.

Number of Adjectives. The number of adjectives produced by a child in a naturalistic setting did not predict their overall lexical diversity in the same setting, indicating that children who use more adjectives do not necessarily have larger more diverse vocabularies as hypothesised, and nor are adjectives associated with the vocabulary gap. It should be acknowledged however that adjectives did form part of the stepwise regression model as a predictor variable, which was not the case for the other non-significant variables discussed above. As further variables were added however, it became non-significant, presumably due to sharing variance with other variables in the model, but not predicting significant unique variance beyond that. Nouns and function words, the other two lexical categories measured, were removed from the analysis due to multi-collinearity between these and other lexical categories.

It was hypothesised that adjectives would form part of the vocabulary gap due to them tending to be more difficult to learn, and therefore acquired later (Booth & Waxman, 2009). This is possibly due to the conceptual complexities of working out which property the adjective is denoting (Tribushinina et al., 2015), but also due to the relative lack of adjectives in parental input (Sandhofer et al., 2000), with both Noccetti et al. (2015) and Tribushinina et al. (2013) finding high levels of correspondence between maternal adjective input and child adjective output. Very tentatively, it is possible that in this particular study, parental use of adjectives was very high (contrary to that suggested by Sandhofer et al., 2000), and this could be due to the specific context from which the data was collected: parents being knowingly filmed playing with their child amongst brightly coloured and age-appropriate toys may influence their language into becoming more descriptive than usual. Therefore, one might conjecture, if maternal input on adjectives corresponds with child output, these particular children may have used a relatively high number of adjectives on this occasion, in reflection of the input they received, and therefore adjectives could not conceivably form part of the vocabulary gap here, as they were used consistently by all children in the sample. Across the general population however, and in alternative contexts, adjectives might be important predictors of vocabulary.

Demographic Risk Factors associated with Children's Vocabulary Outcomes (Research Objective 2)

Research Objective 2 was to investigate the most salient demographic risk factors associated with smaller vocabularies at pre-school age. The reason for this, as noted in Chapter 2, was that demographic risk factors, in the absence of language or behaviour-related red flags, might be useful tools to identify children who might need extra support with their language development. Despite many variables having been associated with language outcomes within existing research, there is some uncertainty around the extent of their influence and relative importance.

Paternal Age and Education

Analysis of the LuCiD data has shown that both a father's age and his level of education is predictive of a child's overall lexical diversity, indicating that older fathers and fathers with a higher level of education are associated with children with larger more diverse vocabularies, and conversely younger fathers and those with lower levels of education with smaller, less diverse vocabularies. This was only true when VOCD (as opposed to NDW) was the outcome measure. These results were corroborated by similar analysis of the MCS dataset, but only for paternal education, not age. Here, paternal education was a significant predictor of a child's BAS score, with a higher level of education associated with higher scores.

The Role of Fathers in General. Until relatively recently, a lack of research (Pleck, 1997) meant that the role of fathers in their child's early development was somewhat unclear. A plethora of recent studies however are generally consistent with the findings of the current study: fathers, either by their characteristics, resources or behaviour, play a significant role in children's developmental outcomes, including language (Tamis-LeMonda et al., 2004; Pancsofar & Vernon-Feagans, 2010; Conica et al., 2020), although their role may vary at different stages of development (Cabrera et al., 2007) and is embedded in the 'larger ecology' of the family, including the family's resources (Lamb, 2004). It is possible that the importance of fathers' interactions with their children has grown over recent years due to altered demographic trends. A greater proportion of mothers in work mean shared childcare responsibilities have become more common (Burgess & Goldman, 2022), and fathers more 'present' than in previous generations. This is particularly true of the LuCiD dataset, where fathers had more caregiving responsibilities than the average UK father (Poulter, 2010), although as discussed in more detail below, this was not a significant predictor of either vocabulary measure.

Father's Education. A number of studies have found a fathers' education to be positively related to their children's language outcomes. In Cabrera et al. (2007), analysis of observational data showed that fathers who engaged positively with their children (scoring high on supportiveness and low on intrusiveness) had significant positive effects on children's cognition and language at ages 2 and 3 years, and that the higher their level of education, the more likely they were to engage in this sort of behaviour, as well as being more motivated and more aware of a child's developmental needs. Indeed, a father's attitude towards learning and their own role in it are likely to be important (Weisleder & Fernald 2013), as more educated fathers are more likely to provide stimulating activities that facilitate talk (Jeong et al., 2017), as well as have a more diverse vocabulary themselves (Malin et al., 2012).

A passive gene-environment correlation is also likely to be a factor in explaining how a father's education may affect his child's language skills (Puglisi et al., 2017). More highly educated fathers may have a higher IQ and more efficient cognitive skills such as attention span, processing speed and short-term memory – genetic material that will be passed onto his own children. He is also more likely to have an educated partner (Watkins & Meredith, 1981) and, as discussed in the preceding paragraphs, value the beliefs, attitude and behaviours that facilitate a child's vocabulary learning.

Finally, it is worth noting that many years of formal education are not a mandatory requirement for creating an optimal language learning environment (e.g., motivation, attitude, diverse vocabulary, awareness of developmental needs). Indeed, ability to create this environment includes practical skills such as incorporating stories, songs, nursery rhymes, trips out and shared activities into daily routines

(Moss & Washbrook, 2016), alongside a warm and sensitive parenting style. With this in mind, perhaps one of the most important explanations for how paternal education can affect a child's vocabulary is via its links with poverty and income, in that those with lower levels of education are less likely to have secure and flexible jobs, and might be experiencing the adverse effects of poverty. According to the Family Stress Model and Parental Investment Model (Perkins et al., 2013) the emotional stress and lack of time and energy resulting from poverty can lead to a more authoritarian parenting style with less opportunities for interaction and nurture, all of which can influence the verbal input children receive (Leigh et al., 2011; Schwab & Lew-Williams, 2016).

Paternal Age. To reiterate, analysis of the LuCiD data indicated that older fathers were associated with children with more diverse vocabularies. Plausible explanations for this are that older fathers have had more opportunities to gain valuable experience that could be utilised in child rearing (Mare & Tzeng, 1989), and older parents are less likely to have had their education disrupted (Duncan et al., 2018) and are more likely to have higher income jobs with some elements of flexibility (Jafee et al., 2001). They are therefore less likely to feel the stressful effects of poverty, nor are their children likely to experience fewer nurturing interactions because of this stress, as per the Family Stress Model and Parental Investment Model noted above (Perkins et al., 2013).

On the other hand, this finding was not reflected by the MCS data, where paternal age was not predictive of BAS scores. One possible explanation for this contrasting result could lie in the different characteristics of the datasets, and in particular the 13-year gap between when the MCS and LuCiD data were collected (2004-2017). A father's age may have been less influential on children's language development in 2004 due to subtle demographic differences at that time: fathers tended to have less caregiving responsibilities than they do now (Burgess & Goldman, 2022) and therefore less interactions with their children, making any experience they may have gained through age less influential. In contrast, the LuCiD fathers had a higher-than-average level of caregiving responsibilities, making it more likely that their characteristics, including age, would have an influence on their child, although again the binary nature of the response for paternal caregiving make any conclusions highly speculative. The MCS sample also had a considerably lower median income than the LuCiD sample, as well as a lower mean age, suggesting that the links between age, income and the stressful effects of poverty may be less pronounced here – as even fathers who are older are not necessarily earning enough to allow them the flexibility and security of those in the LuCiD dataset.

In summary then, results from the current study indicate that a father's age may be important for his child's language development, but the extent of this influence may depend on the characteristics of the sample, particularly income, job flexibility and caregiving responsibilities. Existing literature on the effects of paternal age on children's language outcomes is similarly inconclusive, although it appears likely that a U-shaped relationship exists, with very young fathers and very old fathers associated with the biggest risks for children's development (Malaspina et al., 2005), but again this may be confounded by other factors, such as SES and maternal age, and Gajos and Beaver (2017) found this only to be true for male offspring.

Summary: Paternal Age and Education. This study has found that both a father's age and his level of education is predictive of a child's overall lexical diversity, indicating that older fathers and fathers with a higher level of education are associated with children with larger more diverse vocabularies, but only when VOCD was the outcome measure. These results were partially corroborated by similar

analysis of the MCS dataset, but only for paternal education, with differences between the samples discussed as a possible explanation for this. Recent existing research appears to be generally consistent with the finding that fathers play an important role in their child's development, including language, and this role is probably growing in influence due to shifting demographic trends around childcare. Possible explanations for the influence of paternal education are that more educated fathers may be more likely to have diverse vocabularies themselves, as well as being more aware of children's developmental needs – although it was noted that formal education is not a mandatory requirement for such things. Plausible explanations for the influence of paternal age are that older fathers have had more opportunities to gain valuable experience about child rearing, are less likely to have had their education disrupted, and thus are more likely to have higher income and flexible jobs. It is this link with income and poverty through which both paternal education and age might have the most influence on children's language outcomes. Those with lower levels of education (possibly because they became a parent at a younger age) are less likely to have secure and flexible jobs, and might be experiencing the adverse effects of poverty which may inadvertently affect the opportunities they can provide for verbal interactions with their children.

Gender

Analysis of the LuCiD data has shown that gender is predictive of a child's overall lexical diversity, with girls more likely than boys to have larger, more diverse vocabularies, but only when NDW was the outcome measure (as opposed to VOCD). This result was supported by analysis of the MCS dataset, where gender was also predictive of language in the form of BAS scores, with boys tending to score lower.

Girls outperforming boys in language is generally supported by the existing literature (e.g., Dunatchik et al., 2018; Law et al., 2019), although not unfailingly (Rowe, 2012). Moss and Washbrook (2016) used MCS and NPD data to disentangle the gender gap at age 5 and 11, suggesting that whilst parental reading and trips to the library were equal between genders, there were other subtle differences in early interactions reported by caregivers that may account for differences in the development of language skills, including the fact that girls were more likely to spend time drawing, painting, and singing songs and nursery rhymes with their mothers, whilst boys were more likely to take part in sporting activities. Mothers also reported that they were more likely to teach their daughters letters and sounds than their sons, but this bias did not apply to numbers. These subtle differences in early interactions go some way in explaining the gender differences in language development, but they do not illuminate whether these differences are driven by parental beliefs on gender roles, or the child's own behaviour (that may be biologically driven), or indeed a combination of the two.

Another possible explanation for the differences in how boys and girls acquire language could be the link between socio-emotional development and language acquisition, as supported by both the social-pragmatic (Akhtar & Tomasello, 1996) and the Emergentist Coalition (Hollich et al., 2000) models of word learning. Slot et al. (2020) found that there was a bi-directional influence between children's socio-emotional skills and language skills, with language skills having a stronger predictive influence for boys' socio-emotional development than for girls, and boys tending to have poorer socio-emotional skills than girls (Matthews et al., 2009). Bierman et al. (2009) also found girls to exhibit more prosocial behaviours in a preschool environment, however in this case, despite the difference in

behaviours, there were no associated differences in cognitive readiness scores (some of which incorporated elements of early literacy), suggesting that the explanation of girls' initial language advantage via the link with socio-emotional aspects of development needs further clarification.

In summary, the current results broadly support the notion that gender influences language outcomes in early childhood. This could be due to differences in early parental input, or to more biologically driven differences in behaviour and preferences, or probably a combination of both. Another possible explanation for the gender gap in this context is the link between language skills and social skills, and the different ways these skills are thought to develop in boys and girls, although this explanation has been questioned and thus requires further clarification.

Dummy Usage

Analysis of the LuCiD data has shown a negative relationship between increased dummy usage and a child's vocabulary, although only of borderline significance, and only when VOCD is the outcome measure. The MCS has no comparable data on dummy usage. Existing literature appears to suggest a similar but not necessarily significant relationship (Barbosa et al., 2009; La Prairie, 2010), with some studies indicating no relationship at all (Shotts et al., 2008). Importantly, however, these studies all use simplistic measures of dummy usage that do not account for frequency of use during the day. This is directly comparable to the current study, whereby responses were categorical in nature and only accounted for whether a dummy had been used at all in the past or currently, or had never been used – not how frequently it was used and at what time of day or night. In Strutt et al. (2021) duration of dummy usage (in months), frequency of daily use and times of day were recorded, with results indicating that only prolonged daily use had a possible impact on speech, in that it was associated with an increase in the number of atypical phonological errors. This was likely because of both articulation issues due to dental malformations (Boshart, 2001), but also due to the reduced opportunities for babbling and early word practise (Burr et al., 2021), with the latter only being an issue when a dummy was in use during the day. Children who speak less tend to elicit less caregiver talk, therefore a reciprocal spiral ensues with less quantity and quality of input and smaller vocabularies (Huttenlocher et al., 2010).

Non-significant Variables

The LuCiD data indicates that maternal education, maternal age, the number of children living in a household and income were not predictive of children's overall lexical diversity as measured in a naturalistic setting, which contrasts with MCS data measuring the effect of similar variables on children's standardised vocabulary scores. Paternal caregiving, only measured for the LuCiD cohort, was similarly non-significant. Each variable will be considered briefly in turn.

Maternal Age and Maternal Education

Neither Maternal Age or Maternal Education were found to be predictive of a child's overall lexical diversity, or there was insufficient evidence to say one way or another. This unexpected result is difficult to reconcile with both the MCS result, and the multitude of studies that have said otherwise (e.g., McGillion et al., 2017; Moss & Washbrook, 2016). It also appears at odds with the significant result found for paternal education, as the quality of mothers' and fathers' parenting has been shown

to be similar (Cabrera et al., 2007; Rowe et al., 2004), and to have a similar effect on their children (Moss & Washbrook, 2016), with parental education of one parent predicting the other parent's efforts (Jeong et al., 2017). For Maternal Age, the primary mechanism through which older mums were thought to benefit their children was through more years of education completed, as well as improved mental health (Duncan et al., 2018), greater emotional support (Moore & Brooks-Gunn 2002) and knowledge about developmental milestones (Fergusson & Woodward, 1999).

One possible explanation for this anomaly is the LuCiD sample itself. Mothers were, on average, more highly educated than fathers, with most of the data (85%) concentrated in the degree/postgraduate response category. This is relevant because it suggests that these women had established and well-paid careers at the time of their child's birth, making it conceivable to speculate that they continued with their careers, at least on a part-time basis, and shared caregiving responsibilities – therefore their children's language development may not have benefited from the extra time that they had spent in education, potentially gaining a broader vocabulary and additional knowledge about developmental milestones - as it might have otherwise done. This explanation fits with the detail that LuCiD fathers appear to have more caregiving responsibilities than the average UK father (Burgess & Goldman, 2022), although it should be noted that the binary nature of the response for paternal caregiving make any inferences somewhat speculative. On the other hand, the MCS mothers had more typical levels of education, and in 2004, would have been more likely to have a greater proportion of childcare responsibilities than women in 2017, and a greater influence on their child's outcomes.

In summary then, maternal age and education may be a risk factor for language delay, but this seems to depend on the characteristics of the sample, as with paternal age. Existing literature on the effects of maternal age on children's language outcomes appears to support the idea that older mothers have a positive influence (Duncan et al., 2018), although, as with paternal age, it appears likely that a U-shaped relationship exists, with younger mothers (those under 20 years old) and older mothers (those over 45 years old) associated with the biggest risks for children's development (Dunatchik et al., 2018).

Income

Income was not predictive of a child's overall lexical diversity for the LuCiD data, a result that is somewhat at odds with both the bulk of existing literature, which tends to show a significant relationship between SES and children's language outcomes (Hart & Risley, 1995; Hoff, 2003; Rowe 2017), and also the MCS data, where income was predictive of vocabulary scores. Yet both the assumption that poverty/lower SES puts children at risk of lower achievement and the assumption that poorer outcomes and using income/SES as a blunt tool for intervention have also been questioned in Chapter 2. The role of poverty/SES is likely to be an exacerbating one, but as noted back in Chapter 2, stress and time constraints (which may adversely affect parent-child interactions) exist across all backgrounds, and poverty or particular social backgrounds certainly do not guarantee adverse outcomes. Indeed, creating an optimal language learning environment for children (e.g., ensuring that there is plenty of quality verbal input, shared reading, songs and nursery rhymes) is not dependent on income per se, but can sometimes be affected by factors related to income, such as the stress of poverty or noise levels affecting the ability of parents to provide optimal input. It also difficult to disentangle the effects of income from related factors such as education and attitudes about learning

and development. Parents in the LuCiD cohort volunteered their time for the purpose of educational research, suggesting that all of them, regardless of income level, had an interested and positive attitude towards children's development. With these characteristics in place, income (as a proxy measure for other things) becomes less relevant.

Number of Siblings/Children in Household

The number of siblings a child has was not predictive of their overall lexical diversity in the LuCiD data, or there was insufficient evidence to show such an effect one way or another. This result was contradicted by the MCS data, which did show a significant relationship between the number of children in a household and children's vocabulary scores, a result that is generally supported by the wider literature, in terms of vocabulary or other language-related skills (Laing & Bergelson, 2024; Prime et al., 2014; Moss & Washbrook, 2016; Muñez et al., 2022).

Characteristics of the LuCiD sample may shed some light on these contradictory findings. Many (almost 70%) of the LuCiD sample were firstborns, which may put them at a slight advantage for initial productive vocabulary (Hoff-Ginsberg, 1998), and by 36 months over half remained only children, indicating that number of siblings would have had zero effect on a substantial proportion of the sample. Of those with siblings, only seven participants had two or more siblings, and only two of these seven had three or more, which Laing and Bergelson (2024) found to be the important threshold for negative influence. Any negative repercussions arising from multiple siblings (e.g., lack of space and fewer resources per child) could also conceivably be mitigated by the higher-than-average incomes of the LuCiD families.

Paternal Caregiving Responsibilities

The level of paternal caregiving a child received was not predictive of their overall lexical diversity in the LuCiD data, with no known comparable data on paternal caregiving responsibilities for the MCS dataset. Fathers in the LuCiD dataset did appear to have more caregiving responsibilities than the average UK father (Burgess & Goldman, 2022), although it should be noted that the binary nature of the response for paternal caregiving, and the varying and limited ways in which data about paternal childcare are collected, make this a highly tentative assertion, as previously noted. For example, Poulter (2010) notes that fathers partaking in regular and shared childcare amounts to 24% of all fathers, and Maisey et al. (2013, cited in Burgess & Goldman, 2022) found 53% of cohabiting fathers look after their infant at least several times a week, with 22% doing so every day. Whilst these are both considerably less than the 78% of fathers who answered yes in the LUCID study, 'yes' in the latter could indicate a range of different things in terms of frequency, quantity and range of activities. Both comparison sources are also over ten years old, which means that the proportion of fathers involved in childcare could have increased even further in recent years. Burgess and Goldman (2022) note that collecting data on the fathers' childcare involvement has become more difficult recently, due to family questionnaires not distinguishing between maternal and paternal contributions.

Existing literature has suggested that fathers have an important role to play in their child's development, both directly and indirectly (Tamis LeMonda et al., 2004), via their child's education (Pancsofar & Vernon-Feagans, 2010), sensitive and stimulating parenting (Tamis LeMonda et al., 2004) and repetition (Conica et al., 2020). These effects hold even when controlling for maternal input

(Pancsofar & Vernon-Feagans, 2010), and have been found even in the absence of a maternal effect (Conica et al., 2020). This evidence supports the idea that fathers play an important role in their children's development, however there is nothing here to state that children' language outcomes will improve with every extra hour of involvement from a father, but propose the importance of at least some quality time spent together – fathers who spend no or very little time with their children will have no opportunity to demonstrate the characteristics which the studies above have shown to be important. Thus, the finding of the current study that paternal caregiving is not predictive does not indicate that fathers' involvement is not important for language outcomes, only that the measure is limited in what it could tell us, and future research should measure both quantity and quality of father involvement where possible as 'good child outcomes should be better predicted by the combination [of both] than by either dimension alone' (Pleck, 1997, p.69). Future research is also needed regarding the important and complementary role that fathers can play, as suggested by (Tamis LeMonda et al, 2004), and specifically how this interrelates with a mother's role. Indeed, perhaps there is an optimal balance of shared childcare that exists whereby each parent provides particular elements of input, resulting in the ideal combination for children's outcomes.

Non-significant Variables: Summary

The variables that were found to be non-significant by the LuCiD analysis were often at odds with both the MCS analysis (intended as a triangulation point) and existing wider research. For the most part, the discrepancies can be tentatively explained via the specific characteristics of the LuCiD sample, including highly-educated mothers and relatively high numbers of only children, as well as the complexities and uncertainties involved in the influence of both income and a father's childcare involvement on children's language outcomes. Employment status, whilst only included as a supplementary point of interest, and not intended as a triangulation point with the LuCiD data (which contained no comparable variable), was also found to be non-significant in terms of its relationship with children's language outcomes, which is somewhat contrary to existing evidence which has linked unemployment to children's achievement outcomes in general (e.g., Siddiqi et al., 2007; Jenson, 2023). This result could be explained by either the way the variable was measured (as noted in Chapter 4, it was difficult to re-code the varying levels of the categorical responses in a numerical order, particularly when the status of one parent was unknown, or the family was single-parent), and the fact that the MCS dataset had very low levels of overall unemployment - both parents not in work - which is the combination thought to be most influential on children's outcomes.

General Discussion

The aim of this study was to investigate the expressive vocabulary differences evident in pre-school children in a natural setting (using the LuCiD dataset), with a view to ascertaining whether vocabulary (measured in terms of lexical diversity) could be predicted by particular linguistic and demographic measures, with demographic results then triangulated against the MCS, a much larger dataset that used similar demographic variables and measured language using the BAS. There was a complex profile of results, with only verbs emerging as a clear predictor for both measures of lexical diversity. Additional tentative results show a possible role for rare/less frequent words, gender (as predictors of VOCD), and paternal age and education (as predictors of NDW). A wide range of demographic variables predicted BAS scores in the much larger MCS dataset, including the number of children in a household, gender, income, maternal age and education and paternal education. The

first part of this chapter has discussed each variable in turn, with variables grouped under the two core objectives for Phase One. In the next part of the chapter, a general discussion will consider the different results between vocabulary measures and datasets, and the challenges and limitations associated with using naturalistic language samples and AofA research to make predictions. It will also discuss the differing amounts of variance explained by the statistical models, the implications of using simplistic measures of vocabulary and of the chosen method of statistical analysis, before concluding and considering implications for future research.

Reconciling the Differences between NDW and VOCD as Measures of Lexical Diversity

The current study included both VOCD and NDW as outcome measures of vocabulary diversity in separate models with the same predictor variables, with the expectation that each set of results would be similar and validate the other set, largely due to the finding that different measures of vocabulary tend to be highly correlated, albeit amongst a younger age-group (Bornstein et al., 1998). With the exception of verbs, this expectation was not realised, and the results obtained from the VOCD and NDW models were inconsistent with each other, possibly explained by the different methods used to calculate each measure, which will be described below.

VOCD is based on the probability of new vocabulary being introduced as the speech sample increases in length, accounting for more than just vocabulary range and size, but also the way vocabulary is deployed beyond a word's initial use (Malvern et al., 2004). It is often presented as a gold standard measure because it is not a function of the number of words in the sample, uses all data available, and is more informative due to representing how type-token ratio (TTR) varies over a range of token sizes. It is calculated by pitting the TTR against the token curve for the whole transcript (see p.111 of CLAN manual linked from the CHILDES database, MacWhinney, 2000). Controlling for transcript size is important, as larger samples with more tokens tend to yield smaller values for TTR, and vice versa (Jones & Rowland, 2017) as each time a new word is added, this increases the token count, but only increases the type count if the word has not been used before, hence TTR will inevitably fall as sample size increases (Durán et al., 2004). It should be acknowledged however that VOCD is not universally accepted as the gold standard for measuring lexical diversity, and its ability to account for the number of words in a speech sample has been questioned across different transcript lengths (DeBoer, 2014; McCarthy & Jarvis, 2007).

VOCD tends to be a more conservative estimate than NDW due to it being computed based on the number of lemmas (with different versions of the same root word counted as one lemma). MacWinney (2023) asserts that the lemma should be the superior unit of analysis, as lexical diversity measures should only be concerned with a child's knowledge of different dictionary words, not their ability to use different tenses and plurals. Yet this is not an uncontentious statement, and is somewhat at odds with the pragmatic approach intended for the current study. For example, the words *fish* and *fishing* are from the same root word, and would be treated as the same word type under the VOCD method, and yet to a child, not only do they sound different, but they can also be employed in entirely different contexts with different meanings – suggesting that it is equally justifiable to count them as different types, as per the NDW system. In other words, if a parent, carer, or child is likely to categorise *fish* and *fishing* as different (albeit similar) words, then only counting different lemmas as different word types may misrepresent this information.

NDW uses a broader classification system that is not based on lemmas, but rather allows for different versions of words (e.g., *do*, *don't*, *does* and *fish*, *fishing*) to be recognised as different word types. Despite VOCD being widely used as a measure of vocabulary diversity, it is not universally used, and studies with large samples often count every unique word as a word type (Hart & Risley, 2003; van Heuven et al., 2014), including proper nouns (Jones et al., 2021). In this particular study, NDW was also standardised over time (as all the video samples were the same length), which should mean that the measure is not confounded by sample size (Snow et al., 1996), yet Duran et al. (2004) argue that even then the NDW method is still confounded by volubility and frequency, in that it measures how talkative someone is during the time window given. It is likely that the circumstances of naturally occurring speech samples may encourage some children to speak the words they know more than others, but this is a general limitation of collecting observational data, rather than one associated specifically with using NDW as a measure of vocabulary.

The preceding paragraphs offer some justification as to why both VOCD and NDW were used as measures of vocabulary diversity in the current study: VOCD because it accounts for token size, repetition and uses all data available, and NDW because it aligns with the study's pragmatic approach, and standardises the speech sample via an identical time window. On the other hand, they are less fruitful in explaining why the two measures, with the exception of verbs, produced different results with the same set of predictor variables. Methodological differences in how each measure is calculated, particularly VOCD's use of lemmas and its focus on accounting for volubility and deployment (beyond vocabulary range) are likely to be behind such differences, but clearly further research is required to clarify these issues, possibly with larger speech samples and more tightly controlled input from adults.

Discrepancies between the Amount of Variance Explained by the Different Models

Three models were tested in the main analysis with a view to simultaneously addressing both Research Objective 1 (linguistic characteristics) and Research Objective 2 (demographic risk factors), with Models 1 and 2 using the LuCiD dataset and two different measures of vocabulary (VOCD and NDW, respectively), and Model 3 using the MCS dataset, where BAS (naming vocabulary subtest) was the outcome measure. Two further models (Models 4 and 5) were subsequently tested with the intention of comparing how different children used the same verb within an utterance, thus shedding light on whether children who used more verbs tended to employ additional word types around those verbs. Despite a significant correlation between VOCD and NDW (see Table 2, Chapter 5), the differences between the two vocabulary measures have been discussed in the preceding subsection and could partially explain some of the differences in variance explained by Models 1 and 2, as well as Model 3, (MCS data), which had a lower level of variance explained (7%) than both the LuCiD models.

Further consideration is required regarding the large discrepancy in variance explained by Models 1 and 2 based on the LuCiD data (37% for the VOCD model; 92% for the NDW model). One plausible explanation for this possible discrepancy is the closely-linked relationship between some of the predictor variables and the dependent variable in the NDW model, and the way in which these variables were operationalised via raw counts. For the NDW model, number of verbs (for example) is directly linked to the overall number of words, in that a 1 unit rise in verbs will result in a 1 unit rise in NDW – although NDW is not solely influenced by verbs, but also other variables in the model, be they word type counts or other types of variable. As noted in the subsection above and within Chapter 4,

the NDW measure is a simplistic and direct reflection of the diversity of word types used, and this corresponds with the amount of variance explained by the model in which it was the outcome variable. VOCD, on the other hand, is a more complex measure which involves the use of lemmas (thereby discounting different forms of words as different types), as well as accounting for sample size using probability. A possible influence of these methodological steps in calculating VOCD could be a dilution of its relationship with diversity of word types used – unlike the more direct relationship seen with NDW – and thus a lesser level of variance explained by the model in which it was the outcome variable.

In summary then, degrees of similarity between some of the predictor variables and dependent variable in the NDW model are acknowledged, in that they are not completely independent, and this is a possible explanation for high levels of variance explained by the model. Additional subsequent analysis carried out by Models 4 and 5 was intended to address the inherent limitation of any results obtained here (i.e., perhaps verbs were predictive of vocabulary size simply because one more verb = one more word type, as described above), and to ascertain whether verbs were driving vocabulary across multiple word categories (which was tentatively affirmed).

Using Naturalistic Language Samples: Challenges and Limitations

Using a corpus of children's talk to look for patterns in their speech was justified as a methodological approach in Chapter 4, with the rationale including high ecological validity and a bank of data with no ceiling effect. Caution, however, should be employed when making comparisons between research that is based on naturalistic language data, versus other ways of studying and measuring children's language, be they focused on expressive or receptive vocabulary, and utilise parental report methods or standardised assessments carried out by a trained practitioner (despite evidence that language data of all types is likely to be positively related, Bornstein et al., 1998). As discussed in Chapter 1, whilst expressive and receptive vocabulary development are inextricably linked, they are also disassociated to some degree, and thus a child's comprehension of language, measured via parent report or otherwise, may not be reflected by their production knowledge or indeed performance, particularly where there is only a ten-minute recording window. As with all methods for measuring language, 'natural' language samples are still subject to some level of bias, and cannot be said to occur 'spontaneously, without any external manipulation' (Lust & Blume, 2016, p.74-5). Indeed, it is clear that the words a child produces during such a recording will depend on many factors, including their own personalities, the behaviours of the adults present, and the range of toys and books available. As noted by Ambridge and Rowland (2013), the words children are able to express (their linguistic knowledge) can only be inferred from their observed behaviour (their linguistic performance), and the 10-minute recordings analysed here provided only a small snapshot of children's linguistic performance. Even though the current research was interested in capturing all words (as opposed to particular infrequent combinations, for example), there is still a possibility that the snapshot provided could have over or underestimated the patterns seen.

There are also difficulties in comparing results even when language data is of the same naturally-occurring-type and concerned with language production only, as the parameters within which data is collected varies considerably. For example, in Ota and Skarabela (2018) ninety-minute parent-child interactions were collected over a varying number and length of recordings, whereas in Theakston et al. (2001) twice-weekly one-hour sessions every three weeks for one year were used. Obviously, these parameters will depend somewhat on the resources available and the research

questions of interest, but they could also inadvertently influence outcomes too, as recordings that involve children interacting with parents at specified times will almost certainly contain a greater quantity and quality of vocabulary, elicited by an attentive parent who is conscious of the session being recorded.

As noted above, children's linguistic knowledge can only be inferred from their observed linguistic performance (Ambridge & Rowland, 2013), and attributing vocabulary size to such performance necessarily relies on the child being both willing and able to speak the words they know. Whilst this is also true of other methods for measuring language, including standardised assessments, it is more pronounced here, particularly in the playing alone recordings, where children received little prompting or interaction from adults. Scrutiny of the LuCiD videos suggests that some children play quietly (or even silently) because this is what they prefer, and just because an event (in this case speech) does not occur within a given observation window, this does not necessarily suggest that the event cannot happen at all (Lust & Blume, 2016). Indeed, many children were able to demonstrate more convincing language ability during the second video when interacting with a parent. This is important, as whilst VOCD accounts for different token sizes (so can in theory account for varying levels of productive speech across the different videos), it still requires a minimum token size to be able to calculate a reliable score for VOCD. As we know, NDW does not account for token size at all, due to being standardised over time, so a child who decides not to speak, or speaks very little whilst playing on their own, will have likely demonstrated far less of their language ability, and received a lower score for vocabulary diversity.

A child's ability to articulate clearly (and this can also encompass strength of accent) also plays a role in how accurately language knowledge is obtained via transcribing natural language samples. Whilst this problem was somewhat alleviated for the joint parent-child sessions, where parents would often repeat utterances – a useful strategy for clarification and later transcription (Moskowitz, 2000, as cited in Lust & Blume, 2016), or even gesture towards the relevant object making the child's utterance less ambiguous, this was dependent on how each individual parent interacted with their child. The effect of parent-child interaction on language development was not directly addressed by the current research, and yet what the parent did and said clearly played an important role in how the child responded verbally – and will therefore necessarily affect the vocabulary diversity scores they achieved. This also extends to the researcher, who was present in the room to film the sessions, and who may have provided varying levels of input to different children.

A final consideration regarding the use of naturalistic language samples are the environmental conditions within the room where the language samples were obtained, and how these will also influence the outcome of any language measures. For example, the presence of a younger sibling appeared to influence both the participant child's verbal behaviour, and how much (or little) they could be heard and understood, especially if the younger sibling was particularly noisy or demanded a lot of the parent's attention. Auditory distractions have been shown to influence memory performance in adults (Beaman et al., 2014), a finding that could presumably be extrapolated to children in that distractions could affect their ability to retrieve lexical and sub-lexical knowledge. The presence of and preference for different toys also prompted different words to be expressed, as did the tendency (or not) for singing, as, interestingly, even children who appeared to demonstrate limited

speech production were able to reproduce short songs or nursery rhymes verbatim with little prompting.

In summary, the use of data in the form of naturalistic language samples for the current project was a considered and rationalised decision; however, this does not negate the challenges and limitations associated with the approach. Language scores based on naturalistic language samples necessarily only consider productive language, which is only one (albeit inextricably linked) component of language. Such samples also rely on linguistic performance within a particular time window, and this depends on a multitude of factors, including parent interaction, the presence of siblings, the child's personality and the range of toys and books available. Difficulties in comparing findings across studies were also considered, as were the challenges of the transcription process itself (although these were somewhat alleviated by the presence of video footage, as described).

Decisions about Variables and Statistical Analysis

Many studies investigating the influence of linguistic factors on word learning tend to focus on a smaller number of variables, or even a single variable in isolation. Jones and Brandt (2019) argue this to be an unsatisfactory strategy as such variables may prove to be less important when considered alongside a wider range of other potential variables. The current study therefore adopted a similar approach to Braginsky et al. (2016) and Goodman et al. (2008), whereby predictor variables and ways of measuring them were appropriated from multiple sources and chosen with the intention of addressing the research objectives at hand. Whilst this strategy allowed for many variables to be considered simultaneously, it also presented a sizeable number of variables overall. One statistical method that can be employed on such occasions is the stepwise regression, which offers a systematic pruning of available variables, selecting only some for the eventual models based on their predictive power. Stepwise regressions are sometimes criticised regarding their tendency to pick 'nuisance' variables and ignore 'true' ones (Smith, 2018), however this is less likely when the original list of potential predictor variables is supported by theory and empirical study (as here). On the other hand, using a stepwise regression does mean that the exact combination of variables chosen for each model was not an exact replica of any existing study, and many linguistic (and demographic) variables that had been shown to be important by previous research were not included in the eventual models – indeed only five (out of a potential eighteen) were chosen for the VOCD model, and four for the NDW model. This has implications when comparing results.

Reconciling the Differences between LuCiD and MCS datasets

Research Objective 2 was to investigate the salient demographic risk factors associated with smaller vocabularies in pre-school children, using naturalistic corpus data from the LuCiD dataset to achieve this. In conjunction with this, additional analysis was carried out on the MCS sample, intended as a useful triangulation for the LuCiD demographic results. In fact, contradictory results were produced with respect to income, number of children in a household, paternal age, maternal age and maternal education, with the MCS data tending to indicate such demographic variables were significant in predicting children's language outcomes, and the LuCiD data suggesting otherwise (with the exception of paternal age, where the opposite was true). Possible explanations have been considered above in relation to each individual variable, however there are also some general points

to consider with respect to reconciling these different results, as well as some associated limitations with using both datasets to address the research aims.

The MCS sample was considered a suitable data source for several reasons: it had a comparable data point at 36 months, a language measure in the form of the British Ability Scales (BAS), and many of the same or similar measured demographic variables as the LuCiD dataset. However due acknowledgement should be given to the fact that the BAS (naming vocabulary subtest) and analysis of naturalistic language samples are very different methods for measuring vocabulary. The BAS (naming vocabulary subtest) primarily tests a child's knowledge of nouns through picture recognition, with successful performance depending on the child's prior knowledge of nouns. For the LuCiD dataset, both measures of language were lexical diversity measures (VOCD and NDW) based on naturalistic observations, with no standard set of prompts to elicit words known (although adults did make use of verbal prompts – these were not a standard set of prompts and varied from parent to parent). It is conceivable, then, that the discrepancies between the MCS and LuCiD results could be explained by the different ways each dataset measured language, which adds doubt to the idea that different measures of language tend to positively correlate (Bornstein et al., 1998). Furthermore, not only was the outcome variable measured differently, so too were the predictor variables, which again could have influenced the outcome. For example, income was organised into six different bands for the MCS data, but only four for the LuCiD data, with parental age data also organised in a different way by each dataset.

Specific characteristics of the LuCiD sample are likely to have influenced the results, as discussed above in relation to individual variables. This is particularly apparent when 50% of the speech sample for each child was taken whilst interacting with a parent, with different levels of input having the potential to elicit different levels of speech from the child. The LuCiD sample is characterised by higher-than-average incomes, levels of education, and possibly paternal caregiving, whereas the MCS sample is more representative of the UK average at that time. These differences are likely exaggerated by subtle demographic changes over the 13-year time difference between samples, such as an upward trend in parental age (Office for National Statistics, 2017), less children per family, as measured by average total fertility rates (Office for National Statistics, 2022) and a reduction in absolute child poverty levels (as judged by wealth inequality, which is much lower than it has been in previous years, Szreter, 2021, although short-term trends may be susceptible to fluctuation) - and could also have stemmed from the samples being recruited from culturally and geographically different areas: Bristol and the North West.

Finally, and possibly most importantly, there is a huge difference in sample sizes between the MCS and LuCiD datasets, with the MCS containing data on 7652 children (once missing data had been removed) and LuCiD just 73 children. It is well established that sample sizes can be both too small and too large, with the former risking not finding true differences that exist in the population, and the latter risking over-powering differences that in real life are irrelevant (Faber & Fonseca, 2014). A larger proportion of predictors in the MCS model were found to be significant (6 out of 8) when compared to the LuCiD models (3 out of 18 for the VOCD model, and 4 out of 18 for the NDW model), and the effect sizes for the latter tended to be of the medium – large variety, whereas those for the MCS model tended to be small or occasionally medium, suggesting they may be less practically relevant. A further

explanation for the differences encountered could also be the use of different analytical methods (a multiple regression was used for the MCS data as opposed to a stepwise regression for the LuCiD data).

In summary, there were multiple discrepancies between the results provided from each dataset, despite the variables being similar in nature. Some of the general reasons for this are considered above, including the differences in the outcome variables used to measure vocabulary, the differing characteristics of the samples (e.g., the higher incomes and education levels within the LuCiD sample), likely exacerbated by the 13-year gap between the two datasets, and finally the large sample size difference.

Limitations Associated with LuCiD and MCS Datasets

The subsection above attempts to reconcile the differences in the results presented by the two datasets - leading logically on from this is a consideration of some of the limitations presented by the datasets.

The MCS Dataset. The MCS dataset found a number of demographic variables to be predictive of vocabulary in pre-school children, as hypothesised. The dataset was thus a suitable resource for addressing Research Objective 2 (to investigate the demographic risk factors linked to smaller vocabularies). Despite this, when Research Objective 2 is considered alongside Research Objective 1 (exploring the linguistic characteristics of the vocabulary gap), in that the former was intended to be a tool for identifying children with smaller vocabularies, and the latter to identify the word types which may be missing, to inform intervention content – then the dataset is perhaps more limited. This is because, the measure of vocabulary used (BAS, naming vocabulary subtest) was very much focused on children's ability to name objects (nouns), and therefore may have been overly simplistic when considering the vocabulary gap, in that this study (and others, e.g., Hadley et al., 2016) has found other word types beyond nouns to be important, particularly verbs. This may have resulted in children obtaining a low BAS score, and identified as needing extra support, when in fact this low score was only reflective of their noun knowledge, and did not extend to knowledge of other word types. Nonetheless, the BAS is a widely used and valid measure of children's vocabulary, and for the purposes of giving some indication of the important risk factors for children's language, it does a broadly suitable job.

The MCS might also have been limited in its suitability to address its intended Research Objective (2), in that, despite being broadly representative of society at that time, the data is now somewhat outdated at 20 years old. Shifts in demographic trends (as noted above, such as parental age and number of children per family), as well as societal changes not captured within the data, such as parent's and children's increasing use of technology, may have resulted in subtle changes regarding which factors are important for vocabulary learning. For example - whilst extremely tentative - it is possible to conjecture that traditional demographic risk-factors such as income and parental education may have become less influential predictors of vocabulary over time, as they are mitigated by internet connectivity and the availability of free apps and other resources that assist children's learning.

The LuCiD Dataset. For Research Objective 2 (investigating the most salient demographic risk factors for smaller vocabularies in pre-schoolers), the LUCID data, whilst more recent than the MCS, is still

approaching ten years old, and so some of the same arguments regarding age of dataset and increasing influence of technology also apply here, but to a lesser degree. The LUCID sample was also characterised by higher-than-average incomes and education levels, so results of any analysis based upon this dataset may not be fully generalisable to the rest of the population. As noted above, a voluntary sample of parents, particularly for research which involves a relatively substantial time commitment, implies a high level of interest and involvement from parents and their children, as well as the flexibility, time availability and feasibility of being involved. This could also be considered a limitation, in that families who are experiencing the most difficulties, and are therefore not able to feasibly volunteer or commit to such enterprises, are unlikely to be represented here. Such families may be experiencing multiple risk-factors which could influence their children's language outcomes, and not being represented in the LuCiD sample brings into doubt whether the full spectrum of vocabulary ability is represented here (this point is picked up again later in the chapter).

For Research Objective 1 (exploring the linguistic characteristics of the vocabulary gap), the LuCID dataset was chosen due to being the largest and most recent collection of speech data for children aged between 0-5 years (including a 36 month data point) known to the researcher at the time. The 36-month age-group was thought to be particularly suitable for the current investigation, in that it was located at the lower end of the pre-school age range (3-5 years), an age where children are old enough to have vocabulary sizes large enough to analyse and search for patterns, but also young enough to still benefit from a parent-led intervention at home, as was intended for the Action Research cycles in Phase Two. Therefore, LuCID provided a rich set of naturalistic speech data through which patterns associated with larger and smaller vocabularies could be analysed. Limitations associated with the use of naturalistic speech data, particularly regarding children's speech, have already been noted above, and include the possibly confounding influences of context (e.g., toys, siblings), parental input and volubility, suggesting that vocabulary size (as measured here in two ways: VOCD and NDW) might not always be reflective of a child's vocabulary knowledge, depending how much the child was influenced by these factors. For example, if a child was uninterested in the toy selection, distracted by a sibling or overly shy, this may have resulted in a VOCD/NDW score which was not truly reflective of their vocabulary ability.

As hinted at the end of the first paragraph in this subsection, another possible limitation of using the LuCiD dataset to investigate the vocabulary gap (Research Objective 1) is the possibility that the dataset was not able to provide insight into the full spectrum of vocabulary development. Whilst any differences of vocabulary amongst the sample would give some level of understanding into the vocabulary gap, a more thorough exploration would be facilitated by data which contained a range of vocabularies from the very largest to the very smallest. The extent to which this range has been provided depends on the vocabulary measure in question, as detailed in Chapter 5, Table 3, where the range of vocabulary scores, by measure, is considered against norm-referenced samples. This includes an indication of how many children fell below the 10th percentile, and at the other end, above the 90th percentile. The 10th percentile was of particular importance as an indication of the lower end of the vocabulary size range, as this is the threshold which is often (although not always) used to indicate language delay (Perez-Pereira, 2021), or 'slower expressive language development' (Dale & Patterson, 2017, p.3).

Both VOCD and CDI-III scores appear to show a sample which is mostly concentrated below the 50th percentile when compared to norming samples, with few, if any children, scoring above the 90th percentile. The CDI-Lincoln scores on the other hand indicate very much the opposite, with most of the sample scoring above the 50th percentile, and over a quarter scoring above the 90th percentile. It is possible that some of these discrepancies could be resolved by considering the limitations of the norming data (for example the VOCD norms to which scores were compared were based on a sample of only 32 children, with the 10th and 90th percentile scores estimated - see Duran et al., 2004, p.230), and the Lincoln-CDI and CDI-III norms were compared, in the absence of UK norms, to US-based norms (the former to norms based on the original CDI: Words & Sentences). This US norming data may be misleading, as US children tend to score consistently higher than UK children (e.g., Hamilton et al., 2000) – although this would suggest an over-estimation of language delay, rather than the other way around. The fact that 'CDI: Words & Sentences' contains 680 as opposed to 689 words will also have a particularly large influence on the 90th percentile threshold, with many LuCiD children reaching the 90th percentile that perhaps would not have done with a UK-normed reference.

The different sample profiles suggested by the different measures is difficult to reconcile, although somewhat consistent with the result that the different measures of vocabulary were not all significantly correlated with each other, as might have been expected, with moderate-strong correlations only between the two naturalistic measures on the one hand, and the two CDI measures on the other (there was also a weak significant correlation between NDW and CDI-III). The different vocabulary measures may be telling different stories about the data, with the CDI-Lincoln measure indicating a sample weighted towards larger vocabularies, and the VOCD measure suggesting a sample weighted towards smaller vocabularies, with the CDI-III somewhere in the middle. Such ambiguity presents the possibility that the LuCiD dataset may have been limited in its ability to address the aims of this research (Research Objective 1: exploring the linguistic characteristics of the vocabulary gap) which, as noted above, would have been best facilitated by data which contained a range of vocabularies from the very largest to the very smallest. If this range was not present, then the LuCiD sample may have provided sample-specific thresholds for either 'small' or 'large' vocabularies, resulting in the finding of misleading patterns about the differences in children's vocabularies which are not representative of the population in general. Both the CDI-Lincoln and VOCD measures suggest a reduced range of vocabulary sizes within the LuCiD sample, with those ranges concentrated towards the generally larger or generally smaller vocabulary size. The CDI-III indicates a more evenly distributed range (albeit with a much smaller 100-item measure).

In spite of these possible limitations in range (regardless of the vocabulary measure used), it is suggested that some level of insight into the vocabulary gap has been provided nonetheless, with the ambiguity involved in defining language difficulties or delay adding weight to this assertion. In assessing the range of vocabularies evident in the current sample, overall distribution was noted with particular reference to the proportion falling below the 10th percentile (and above the 90th) due to this being an accepted threshold for identifying language delay (Perez-Pereira, 2021; Dale & Patterson, 2017). However, as noted in Chapter 1, the proportion of children with language difficulties is estimated to be far more than 10% (e.g., 20% in Law et al., 2017a, and 40% in Harley et al., 2018), suggesting that analysis of speech samples which are in the 10-40 percentile range could still provide some valid information about the speech of children who, by some methods, would be identified as having language difficulties or delay.

Summary of Limitations associated with the LuCiD and MCS Datasets. Both datasets have been argued to be broadly suitable to address the current research objectives, in terms of being rich sources of data for the age of interest (36 months) and using different measures of vocabulary (the strengths and limitations of which have been discussed), but comparable demographic variables of interest. Despite this suitability, a number of potential limitations with each dataset have also been acknowledged. The MCS, whilst broadly representative of the population at the time, is now 20 years old, and shifting demographic trends will likely have resulted in changes concerning what factors are important for children's language development. This argument also applies in part to the LuCiD dataset, which despite being more recent, is still now over 5 years old, as well as not being completely representative of the general population (e.g., higher incomes and education levels). One of the key questions regarding potential limitations is whether the LuCiD dataset was able to provide the full spectrum of vocabulary sizes, therefore providing a suitably full range from which to be able to address the stated research objective of exploring the vocabulary gap. Whether or not the LuCiD dataset has been able to do this is questionable, with different distributions of data suggested depending on the measure of vocabulary used. Nonetheless, it has been argued here that some level of insight has been provided regardless, as differences are apparent even with a reduced range to explore.

Implications of using Simplistic Measures of Vocabulary

There are varying opinions about when a word should be considered 'known' by a child and therefore forming part of their vocabulary, with some arguing that early expression or comprehension of a word is simply the beginning of the acquisition process (Ambridge & Lieven, 2011). Very few methods of calculating vocabulary, including the lexical diversity measures used here, consider contextual use of vocabulary or depth of vocabulary knowledge, merely that the word has been recognised or used to some degree or another – which could include repeating words just expressed by an adult, or memorised as part of a song or rhyme – this will not necessarily reflect how skilful children are at using the words they know under the time constraints of a back and forth conversation (Hadley et al., 2016). Some acknowledgement of the limitations of vocabulary measures is therefore necessary due to the impact this may have in terms of real-world implications: a child who scores highly on standardised tests of vocabulary may not necessarily be able to use that vocabulary in an age-appropriate way and in varying contexts.

Using Age of Acquisition Research to make Predictions about the Language Gap

The gap that exists between those with the largest and smallest vocabularies has been well documented (Huttenlocher et al., 2010), with our knowledge extending only to the quantitative existence of the gap (Fernald et al., 2013), not its qualitative nature - hence the unique contribution of the current study. Yet the fact that no previous studies had investigated this matter posed a problem with respect to hypothesising about which categories of words may be contained within the gap, and deciding which linguistic variables to include (indeed it was possible that no pattern existed at all).

Age of acquisition (AofA) research provided one avenue from which to gather these predictions, on the premise that words that are shown to be acquired at an earlier age are easier to acquire and associated with less diverse vocabularies (as younger children tend to have smaller vocabularies), and words that are shown to be acquired at a later age are harder to acquire and

associated with more diverse vocabularies (as older children tend to have larger vocabularies). 'Harder' words in the latter category (associated with the older children with larger vocabularies) are therefore more likely to reside within the vocabulary gap.

One limitation of using AofA data to make predictions is that such data assumes that all children are working their way along the same vocabulary acquisition curve, with some children a bit further behind for what we might expect for a child of that age, and are therefore missing particular categories of words. Yet all children following the same trajectory is not necessarily the case. Children with particular conditions are known to acquire vocabulary not just more slowly but also differently to what is considered developmentally normal. Research on late-talkers, for example, suggests that this group may have atypical rather than delayed development (Macroy-Higgins et al., 2016), indicating that they are not subject to the same patterns of acquisition as AofA research suggests, and teaching them words and word categories from within the vocabulary gap so defined may not have the desired effect of overall vocabulary improvements. Whilst the current research was interested in exploring the vocabulary differences between pre-schoolers (the 'gap'), this necessarily involved comparing those with the smallest and largest vocabularies (using the data that were available; see earlier in this chapter for the acknowledged limitations of this approach in terms of the range available). As noted in Chapter 1, it is those children with the smallest vocabularies, with some degree of language delay, that are of particular interest when studying the vocabulary gap; children who do not necessarily have a specific difficulty associated with their language development (such as DLD) but have nonetheless fallen behind their peers and the expected developmental milestones for their age range. Nonetheless, even if there is no specific or diagnosed difficulty, or membership of a specified group (such as late talkers, as above), it is still worth considering the possibility that these children could also have atypical development – and therefore predictions made based on AofA research (alluding to typically developing children) could have influenced both the choice of predictor variables included and the interpretation of unexpected results.

Predictions based on child-directed speak (CDS) were considered as an alternative possibility, in that the characteristics of caregiver's speech tend to be reciprocated in their child's speech, as per constructivist theories of language development (Behrens, 2021), hence by investigating and comparing the speech of parents associated with children at either end of the vocabulary spectrum, it would be possible to make deductions about the vocabulary gap. This option was eventually discounted, as despite much research linking particular aspects of CDS with child language outcomes, very little compared the speech characteristics of different groups of parents as described or associated particular words or word categories used by parents with more diverse vocabularies in children.

In summary, predictor variables were selected and hypotheses constructed for the current research using a similar but not equivalent area of research – AofA - which is based on typically developing children. Predictions could not be gleaned from CDS research due to the lack of available evidence, nor could they be taken from other vocabulary gap studies, as none yet exist. Any results and interpretation of such results should therefore be received with the appropriate caution.

Conclusion and Implications for Future Research

The aim of phase one of this study was to investigate the expressive vocabulary differences evident in pre-school children in a natural setting, with a view to ascertaining whether vocabulary proficiency could be predicted by particular linguistic and demographic measures. The results of various regression analyses supported some of the hypothesised predictions and not others, with the number of verbs a child uses emerging as the most robust result, suggesting that verbs are associated with larger, more diverse vocabularies, and therefore may be the missing link for those children with smaller 'delayed' vocabularies. Other more tentative outcomes have also been considered, as have discrepancies between differing aspects of the results. Possible explanations for the latter have been explored, and have included, for example, differences in sample characteristics and methods of measuring language. Due acknowledgement has also been given to several limitations, including the use of AofA data to make predictions about what sort of words may be present in the vocabulary gap.

As noted in previous chapters, supporting children's language needs remains a political priority due to its established links with a range of quality of life indicators later on, including academic achievement (Norbury et al., 2016), job prospects (Gascoigne & Gross, 2017) social mobility (Law et al. 2017b; Field, 2010) and life expectancy (Gilbert et al., 2018). Despite a substantial proportion of funding being directed towards addressing this challenge, and a generally positive direction of influence for school-based and parent-implemented interventions, effect sizes are often unclear (Walker et al., 2020), and evidence lacking regarding generalised and longer-term effects (Heidlage et al., 2018). In particular, addressing the vocabulary gap may be more effective if interventions focused upon particular words or word categories that have been found to be missing.

With this in mind, and employing the most robust result to emerge from the current analyses, the following chapters of this thesis will explain how verbs have been incorporated into a parent-led intervention with the aim of developing that intervention for future testing on a larger scale. The intervention will be designed specifically for use with pre-school children, with ease of delivery and engagement in mind. Beyond that, future research should aim to test the notion that verbs (and possibly rare words) are an integral feature of early language learning by trialling this intervention and any other that focuses on teaching these specific categories of words. Children should be assessed by a variety of language measures, and at various intervals going forward to ensure that any positive effects have stood the test of time. Ideally, this research should be targeted towards children who have been identified as being particularly at risk of poorer language outcomes. This group may be particularly difficult to recruit, as shown by the gaps in evidence from Walker's systematic survey of language interventions (Walker et al., 2020).

Further research is also needed to clarify some of the more ambiguous results thrown up by the current study, including continuing to investigate the role of different demographic risk factors in poor language outcomes. The current study indicates that gender, paternal age and education are important, but earlier studies, particularly those linked to larger datasets (Dunatchik et al., 2018, Law et al., 2019) have found many more factors to be significant in predicting a child's language ability, including (but certainly not limited to) income (in one guise or another), size of household, maternal education and household noise levels. Furthermore, this range of demographic factors still does not explain all the variance in a child's language ability, and there are likely to be many other as yet unstudied demographic variables that are important, as well as developmental and genetic factors at play. It is clear that this complex interplay of factors affecting children's language development

requires further unravelling, with relative influences made clearer, particularly if demographic risk factors are going to be utilised in a productive way to identify children most at risk of falling behind.

Chapter 7: Reviewing the Intervention Literature and Setting the Context for Phase Two

As noted in Chapter 1, there is a substantial vocabulary gap between pre-schoolers with the largest and smallest vocabularies (Huttenlocher et al., 2010; Fernald et al., 2013). This is important, because those with a smaller range in vocabulary are likely to face difficulties not only in the short term, for example being able to follow instructions and articulate their needs and feelings in a school environment, (Harley et al., 2018), but also in the longer term, with early language skills linked to future academic achievement (Norbury et al., 2016), job prospects (Gascoigne & Gross, 2017) and social mobility (Law et al., 2017b; Field, 2010). These far-reaching repercussions have ramifications at both an individual and societal level, and the DfE has expressed its ongoing financial commitment to improving early years language skills over the years, including a 5-year funding pledge to the Education Endowment Fund in 2018, and a recent post-election commitment from the incoming Education Secretary to fund the Nuffield Early Language Intervention for its fifth year. Many interventions aimed at closing the gap have shown positive results, (e.g., see Dimova et al., 2020, for a review of the Nuffield Early Language Intervention). Yet it remains to be seen whether interventions that include the teaching of specific vocabulary could be made even more effective and 'gap-closing' if they were to focus on the actual vocabulary that resided inside the so-called gap: in other words, the specific words or word types that children with smaller vocabularies appear to be missing. Prior to the current thesis, robust knowledge regarding the qualitative nature of the vocabulary gap was not forthcoming, and hence could not be incorporated into any intervention design.

Beyond specifying the content of interventions, there are further challenges involved in identifying those in most need of support, particularly amidst a context of scarce resources. This identification process is particularly challenging in the pre-school period, where there is limited contact with third parties (such as educational professionals) who are likely to have specialised knowledge useful for flagging potential issues and initiating early intervention. As noted in Chapter 2, many demographic factors have been associated with children's language scores, including SES (see Dunatchik et al., 2018 and Law et al., 2019), and interventions are often targeted using this metric alone, for example via Pupil Premium funding. Despite the robust link between SES and language outcomes, there is also considerable variation in ability and behaviour within SES strata (Weisleder & Fernald, 2013). The underlying factors that that may influence young children's language development (e.g., differing levels of parental verbal input due to stress or illness) could be true of households from any SES background. Investigating a more finely grained picture of risk for children's language across multiple demographic variables is therefore important if such children are going to be identified and provided with timely assistance.

Set against this context of language inequality and uncertain risk factors, the aims of Phase One of this thesis were firstly to determine the qualitative nature of the language gap (in terms of identifying particular word types that may reside within it), and secondly to investigate the most salient risk factors associated with children's vocabulary at 3 years old. Phase Two will now build on these aims and their related findings to develop and refine a children's language intervention resource using Action Research methodology. Specifically, it will incorporate the result pertaining to Objective 1 regarding the importance of verb diversity in children's vocabularies (see Chapter 5 for results and

Chapter 6 for discussion of these results), in that the number of verbs a child uses in a natural setting is consistently predictive (across different measures) of the child's overall vocabulary diversity. This suggests that pre-schoolers with smaller vocabularies may be lacking in verbs, and therefore could benefit from an intervention designed to increase their exposure to verbs, which will then facilitate their acquisition of further words across all word categories. At this exploratory stage, results pertaining to Research Objective 2 regarding the demographic risk factors for language outcomes will not be incorporated into the research design, as this would effectively limit the sample pool available – however future research should be more targeted. The rest of this chapter will review the relevant literature on language interventions, using the information contained therein to formulate the rationale and initial design idea for the intervention to be developed. As in previous chapters, the terms 'parent' and 'caregiver' will be used interchangeably, and in instances where one has been used and not the other, then this has been done in the pursuit of conciseness, and in most cases, both are intended.

Reviewing Language Interventions

An intervention can be defined as an explicit application of therapeutic or educational techniques intended to modify an individual's performance in a designated area (Law et al., 1998). There is no consistent framework for describing interventions, and they can take many forms, including programmes, activities, approaches, resources, training and targets, and can be delivered universally or targeted towards groups or individuals. Thus, when interventions are interpreted broadly, they can include a broad range of approaches which wouldn't always be considered an 'intervention' in the traditional sense. Even when interventions are interpreted more narrowly, there are still a large number to consider. As an illustration of this, and some of the other complexities involved, Roulstone et al. (2012) identified 158 interventions currently in use by Educational Psychologists and Speech and Language Therapists as part of a series of BCRP (Better Communication Research Report) reports in response to The Bercow Review (2008). The authors found many of the interventions (both published and developed by local service providers) to be overlapping in nature, as well as outcome measures inconsistent, making comparisons between or within interventions difficult, with outcome data generally collected at the individual child level rather than the service level. Finally, a number of well-established and widely used interventions were found to have little evidence to support their implementation in practice, suggesting that sometimes pragmatic factors played a role in intervention selection over and above the evidence-base. Numerous charitable and research organisations, including the EEF, National Literacy Trust (NLT), LuCiD, I CAN (Children's Communication Charity), The Communication Trust, National Children's Bureau, Real Action and the Early Intervention Foundation conduct research on interventions in schools and out, and there is a concerted effort to make this evidence base available to practitioners (e.g., the What Works website and the EEF Toolkit). Despite this concerted effort, it is often argued that a disconnect remains between the evidence-based research and the interventions that are in fact taking place (Bercow 10 years on report, 2018).

Despite the large number of interventions available (as noted above by Roulstone et al., 2012), the varying definitions of 'intervention' and discordance in how they might be recorded and evaluated means that ascertaining the exact prevalence levels of interventions currently in use, and whether they match the level of need, has similar difficulties to those involved in ascertaining the exact levels

of children with SLCN, as discussed in Chapter 1. Some interventions, particularly those delivered informally at home, or at the whole-class or school level, may not be officially recorded as an intervention, nor properly evaluated, even though positive effects may result. Large-scale evaluative information about interventions at school level are likely to focus on specific interventions provided to children with more specific language needs (e.g., DLD), and those taking place in the home will likely be focused upon those children who have reached the threshold for accessing additional services, for example Speech and Language therapy – and so not necessarily designed for alleviating the vocabulary gap or language delay in general. On the other hand, many research studies have tested interventions aimed at improving children's language, and specifically vocabulary – and whilst these interventions can sometimes be aimed at or tested on a sample of children with very specific language or developmental needs, often they are aimed at all children who have smaller vocabularies, and thus aim to 'close the gap.' The next part of this chapter will therefore review specific interventions targeting the pre-school age-range and focusing on improving language outcomes, as well as discussing general themes emerging from the literature, and ultimately focusing on parent-facilitated interventions..

Intervention Timing

There appears to be a consensus amongst researchers that early intervention is best (Heymann et al., 2020; Roulstone et al., 2012), partly due to the vocabulary gap emerging early in the second year of life (Weisleder & Fernald, 2013), but also because language trajectories appear to become more stubborn once children reach school age (Law et al., 2017a). This would suggest that, for children who have smaller vocabularies due to less quantity and quality of verbal input (as has been the focus in Phase One of this thesis) then the earlier they receive help, the better. There are of course other factors to consider, and early intervention is not always feasible or desirable. Identifying children in need or at risk of developing language difficulties is not straightforward, and whilst research has associated numerous demographic factors as important predictors of language outcomes (Dunatchik et al., 2018; Law et al., 2019), these are not always agreed upon. Brain plasticity continues as we get older (Banks, 2016), and many language interventions can and do happen beyond this age, often within the school environment (indeed school in itself could be conceived as an intervention) and with reported success (e.g., Dimova et al., 2020) - although in some cases, interventions attempted later on beyond the critical or sensitive period will have only limited impact (Curtiss, 2014). Burnett et al. (2020) argue against early intervention based on deficit views of children's language development and focusing on narrow prescriptive methods, suggesting that such an approach is likely to exacerbate or even create language difficulties that did not yet exist. Although this view contrasts with majority held views on the importance of early intervention, it again encourages caution when considering the timing of an intervention; simply providing a child with a well-designed, evidence-based intervention will not necessarily equal better language outcomes for that child, with the timing of the intervention being amongst the multitude of possible reasons why.

In further support of the importance of intervention timing, and the interplay of timing with other important factors, Adamson et al. (2020) refer to the epigenetic landscape in the form of Waddington's classic image (Waddington, 1957), with a ball (representing the child interacting with the parent) rolling down through the landscape via numerous deep ravines or shallow paths, representing genetic possibilities, with potentially different outcomes dependent on the path

travelled. Depending on how deeply entrenched they are in their existing path, environmental events (such as an intervention) have the potential to shift the ball into another path, thus the potential success of an intervention can be visualised, in that it depends on how susceptible a particular child is to external influence at any given moment. In acknowledging the 'crucial role' of the parent (Adamson et al., 2020, p.63), and how features of the child's world which are removed from their immediate experience (for example the parent's employment, or Government policies on free childcare) may nonetheless still influence them indirectly, the authors here suggest that interventions which work on a multi-layered basis to reflect this multi-layered existence are more likely to be effective (for instance by providing free childcare and travel alongside an intervention programme). Clearly, a balance must be struck between the importance of early intervention, and the feasibility and desirability of early intervention.

Parents as Intervention Facilitators

The crucial role of parents in their child's development, as noted above, is well supported (Adamson et al., 2020; Reece et al., 2010). Both Law et al. (2017a) and Roulstone et al. (2012) insisted that what parents do with their children in terms of the home learning environment which they create was of critical importance, and predicted language development more strongly than their social background. Akin with constructive theories of language development, consistently embedding optimal quantity and quality of input within high-quality verbal interactions is one of the most effective ways to improve children's early language outcomes (Huttenlocher et al., 2010; Schwab & Lew-Williams, 2016; Tamis-Lemonda, 2014). The requirement of consistency suggests that the individual best placed to facilitate language development is somebody who sees the child on a regular basis, and in the birth to pre-school years this is likely to be the parent or primary carer – although of course this is not always the case. As noted in Chapter 1, increasing access to childcare via the Government's 30 hours free childcare pledge means that early years practitioners are also well-placed to facilitate language interventions with pre-school aged children. Indeed, attendance at early childcare settings has been consistently linked with better language outcomes for children (e.g., Davies et al., 2021), particularly for children from families where parents have lower education levels (Becker, 2011), although this is partly dependent on the quality of childcare provided (Melhuish, 2016). Walker et al. (2020) conducted a systematic survey of 190 non-parental intervention studies that addressed communication and language needs in early childcare or school setting, where expressive or receptive language or vocabulary were outcome measures, and found that most studies reported positive impacts for the interventions that were tested, although sample sizes were small and effect sizes uncertain (only 32% of studies reviewed provided effect sizes). Samples were unlikely to include children from a low SES background, or children under 3 years old. One example of an effective childcare setting intervention is the Nuffield Early Language Intervention (NELI), currently funded via the Government's 'catch-up programme,' and involving over 10,000 schools in the first three years of its implementation NELI aims to develop the narrative, listening, vocabulary, phonological awareness and grapheme-phoneme correspondence skills of reception-aged children through daily sessions over a period of 20 or 30 weeks, and has shown to have an impact of between 2-4 months additional progress when using a composite score of language as an outcome measure, which included two vocabulary-specific measures (CELF expressive vocabulary subtest and the RAPT information subtest). Sessions are delivered by specially trained teaching assistants in both small group and individual formats, with each session including a introductory listening activity, revision of vocabulary introduced

in previous session and introduction of new vocabulary, followed by a narrative activity and summary. Sessions are facilitated using comprehensive resource packs including flashcards, badges and puppets.

Despite generally positive evidence for early intervention to improve language in childcare settings, early childcare is not universal, and indeed the additional provision provided by the 30 hours free childcare pledge is only available where both parents are working, and has been criticised due to its hidden costs (Coleman & Cottell, 2019), and lack of available places, particularly good quality ones (Pollard et al., 2023). Despite being relatively low cost, NELI, for example, still requires significant delivery time on the part of early years practitioners, and this is an important practical consideration. Wasik and Hindman (2018) declare that despite the ample knowledge and supporting evidence regarding how children learn words, it is difficult to translate these principles into feasible, replicable classroom practices or interventions, particularly when pre-school classrooms are often constrained by large class sizes and multiple needs, which limit the opportunities for extended adult–child verbal exchanges, including the option for a large proportion of TA time to be spent on a single intervention. This is particularly true of schools in disadvantaged areas, where additional needs tend to be greater and children more adversely affected by increasingly larger class sizes (Zyngier, 2014). Parents, on the other hand, are a cost-effective resource who are likely to be very much invested in their child's best interests, as well as being more likely to be available for 1:1 interactions with their children. For this reason, the remainder of this review will focus on research where the intended facilitator of the intervention is a parent.

Evidence suggests that parent-led interventions can be effective in improving children's language outcomes (Roberts & Kaiser, 2011; Heidlage et al., 2018). In a meta-analysis analysing 25 studies on parent-led language interventions for children with, or at risk of a language difficulty associated with SES, Heidlage et al. (2018) found a modest and significant effect of reading and play interventions, but only for expressive and not receptive vocabulary, and no studies were to report on the extent of generalised or longer-term maintained effects. Reported effect sizes varied based on type or context of the intervention, dosage, and the parent themselves in terms of the skills, knowledge and attitude that they were able to bring to the delivery of the intervention (particularly where implementation fidelity requirements are not particularly rigid but involve some level of interpretation by the parent deliverer). A further complication is added when research looking into the effects of parent-led interventions does not use parents to carry out the intervention. In a systematic review looking specifically at language interventions directed at children from low income families, Greenwood et al. (2020) found that whilst the majority of interventions were intended for parents to implement with their own children, it was often the researchers who carried out the interventions, affecting the ecological validity of the research in that it was unclear whether the same interventions would have had the same effects when delivered by the parents.

Parent-led Interventions

Much of the research on parent-led interventions designed to improve children's language focuses on reading or encouraging verbal interactions (Reece et al., 2010). Indeed, the frequency of joint parent-child activities, including reading, singing and playing games, has been found to have a positive influence on vocabulary learning in the pre-school years (Wood, 2002). In the following section, interventions concerned with talking, responding strategies and play are considered first, followed by reading interventions, and finally the Abecedarian Approach, which incorporates both

elements (although it is important to note that many of the other interventions discussed also include multiple elements, so can't easily be categorised as 'reading' or 'talking').

Talking & Responding Strategy Interventions. Interventions focusing on increasing levels of maternal responsiveness appear to have the largest positive effects on language outcomes, according to a meta-analysis by Roberts and Kaiser (2011). One aspect of responsiveness is contingent talk, whereby a caregiver focuses their verbal interaction upon the infant's current sphere of interest. McGillion et al. (2017) investigated the effect of a low intensity parenting intervention on seventy-three 11-month-olds designed to increase a caregiver's contingent talk by showing them an initial video, then asking them to practise contingent talk for 15 minutes every day for 1 month. The researchers found significant effects in the short term on both levels of caregiver talk and a child's expressive vocabulary (the latter only amongst the low SES cohort, but results were short-lived and did not persist at 24 months). The sample of children used here are younger than the children focused upon in the current thesis (both the LuCiD and MCS children were 36 months old at the time of testing), and so the importance of contingent talk for vocabulary may not extend to older children – indeed some models of word learning (e.g., ECM) suggest attentional cues such as perceptual salience (i.e., what the child is focused on, so relevant to contingent talk) are used more by younger learners, with older learners shifting to linguistic and social cues (Hollich et al., 2000). Nonetheless, it is likely that in practice a variety of different cues are used dependent on the context, and thus that perceptual salience (and with it contingent talk) continues to play a role in children's language learning and expression even as they get older (Abbot-Smith et al., 2017). The short-term effect noted by the authors here (in that the effects on vocabulary did not persist at 24 months) could have been due to the intervention only lasting for 1 month, and there being no guarantee that parents would have continued these contingent talk habits after this period. The authors suggested follow up sessions to mitigate this short-term effect, with age-appropriate skills being taught through various interventions at the appropriate time intervals, such as recasting and expanding children's utterances.

Interventions which use a 'telepractice' format (providing a service from a distance via internet-based technologies have also shown promise as an effective intervention format for language outcomes (Akemoglu et al. 2020). In Feil et al. (2020) an internet adapted version of the PALS (Play and Learn Strategies) programme directed at low-income mothers (with strategies including reading, responding contingently and maintaining attention) was found to be a significant and positive influence on maternal behaviour with a corresponding influence on infant language behaviour, a composite measure that included both expressive and receptive language. In Kachergis et al. (2020) a phone-based video app depicting parent-child activities was also found to be effective, in that whilst the study did not consider children's language outcomes directly, it did find that parents spoke more and made more bids for joint attention in the video group than in the control group, with both behaviours being associated with better vocabulary outcomes for children (Huttenlocher et al., 1991; Cochet and Byrne, 2016). On the other hand, parental lexical diversity, another important predictor of children's vocabulary (Huttenlocher et al., 2010), did not improve after watching the video, possibly due to a tendency for repetitiveness during the prescribed video task. A convenience sample consisting of mostly high SES families also cast doubt on whether results would generalise to different populations, or be maintained after a longer time span. These studies suggest that interventions that make use of technology have the potential to influence parental behaviour and corresponding child outcomes, but

need tweaking to ensure that they are effective beyond their convenience appeal, and are applicable to all.

Reading Interventions. Shared reading commonly features as an intervention context for improving language outcomes. In a survey of 1000 teachers, the majority of teachers testified that 40% of children in their schools lacked the necessary vocabulary to access learning, with a lack of time spent reading specified as a root cause (Harley et al., 2018). Reading has been consistently linked to a range of positive language outcomes for children (e.g., Pillinger & Wood, 2014), and helps to facilitate a verbal environment that is diverse and includes a high proportion of rare words (Beals, 1993). Blewitt et al. (2016) found that when readers employed extra textual strategies, for example contingent verbal interactions about the story's content and new words, this would have a positive effect on promoting children's engagement and vocabulary learning.

One particular form of shared reading which is thought to be particularly beneficial for children's language outcomes is interactive shared reading, which 'typically involves recasts, expansions and open-ended questions' (Noble et al., 2020, p.1879), with examples including pause reading and dialogic reading. Noble et al. (2020) tested 30 to 36-month-olds (n = 150) and their parents over three different interactive reading conditions, two of which involved intervention strategies of training caregivers via video to engage in the particular interactive reading strategy, and the third (an active control group) simply informed parents about making reading a part of daily routines. The researchers hypothesised that parents in the intervention groups would increase their interactive reading behaviours significantly more than those in the control group, with similarly positive gains for children's language outcomes. The researchers were also interested in exploring SES differences, with previous research suggesting that those from a lower SES background may benefit less from interactive reading strategies (Mol et al., 2008). Contrary to predictions, the researchers found that whilst the interventions were effective at changing caregiver reading behaviour, they didn't boost children's language over and above the active control condition, nor was there any effect of SES. A number of explanations were offered for this unexpected result, with the latter two being worthy of some further discussion in the current context.

Firstly, an active (as opposed to passive) control group likely moderated any effect of the interventions on language; secondly, there was a difficulty in knowing whether the interactive reading behaviours taught were consistently adopted by parents in their reading sessions; and thirdly, the 6-week duration of the intervention may have been too short. With respect to duration of the intervention, other researchers have also noted how relatively short periods of intervention may be insufficient to facilitate change, particularly long-term language improvements (e.g., McGillion et al., 2017). Adamson et al. (2020) noted that interventions should be developed with a view to how they will influence subsequent development, not just immediate effects - indeed, long-term sustainable shifts in behaviour were the key to children's language outcomes. The second explanation regarding whether or not parents practised the strategies learned is one of implementation fidelity – and is related to the discussion on intervention duration in terms of its relevance to ensuring useful language-learning strategies are adapted over the long term. Ensuring implementation fidelity is a potential limitation in many parent-led interventions, particularly where fidelity requirements are complex or require high levels of knowledge or skill. Intervention approaches which are simple (yet still effective) are perhaps more likely to be implemented with fidelity by parents, and over a longer,

indefinite period of time. Indeed, those interventions which can work within the context of established routines which have the potential to be adapted are likely to be most effective (Adamson et al., 2020, p.63).

Contrary to the findings in Nobel et al. (2020), Burgoyne et al. (2018) did find a significant effect on language outcomes (including measures of expressive and receptive vocabulary) for PACT (Parents and Children Together), an intervention drawing together elements of interactive reading alongside targeted vocabulary teaching and narrative activities. PACT, as the name implies, was designed to be delivered by parents to their 3-year-olds with the aim of supporting oral language and emerging literacy skills, and consisted of a 30-week programme with daily 20-minute sessions following a consistent structure and routine. Whilst effect sizes were termed 'modest' but 'arguably educationally significant' by the authors (Burgoyne et al., 2018, p.553), on this occasion positive effects were maintained 6 months later. This sustained impact may have been due to the longer 30-week programme, which aligns with comments made above by both Noble et al. (2020) and McGillion et al. (2017) about shorter intervention durations perhaps being insufficient to see and certainly maintain positive effects, but also could have had something to do with the consistent structure and routine incorporated within PACT, which may have facilitated longer-lasting habit shifts. Of further note, the sample of 208 children included in this RCT were from varied SES backgrounds, however a large proportion were from lower SES backgrounds and started off with lower than average language skills at pre-test, suggesting the PACT intervention and specifically interactive reading strategies could be equally effective with children across a range of SES backgrounds – something which had been questioned by earlier research (e.g., Mol et al., 2008).

A book's genre and complexity also appear to be influential with respect to how much extra textual talk takes place during a shared reading activity. Muhinyi et al. (2020) found that when a complex story (involving a false belief central to the plot) was compared to a simple story, more extra textual talk was facilitated, and of a higher quality. This didn't affect the sheer number of questions posed whilst reading, but follow ups on children's responses were more elaborate. Information books appeared to elicit more extra textual talk than story books, despite the latter being the most popular genre within this age group (Curcic, 2023), possibly because children may find the words and pictures contained in information books easier to relate to their own world (depending on the topic involved), although other potential factors influencing these outcomes were also discussed, including the skill and conversational style of the adult involved.

Whilst not technically a 'reading' intervention, the use of visual prompts alone, without accompanying text, have also shown promise as an intervention tool for improving language, albeit for slightly older children aged 4-6 years (de la Rie et al., 2020). Here, the researchers found that using prompting boards, consisting of a complex and detailed illustration, elicited more contributions from children as a proportion of the interaction than the similar shared reading intervention. This was perhaps due to the fact that an absence of script required participants to infer their own narrative for the picture, as well as a watering down of the traditional dominant and passive roles apparent in traditional shared reading, suggesting that it is the interactive nature of learning that results in knowledge being retained - although it should be noted that some reading approaches do attempt to redress this balance to a certain extent (e.g., the interactive reading approaches discussed above, for example Reciprocal Reading). Parental differences in their ability to be imaginative was also acknowledged as a potential

influencer of child outcomes (de la Rie et al., 2020), in that those who found using their imagination more challenging might find more scripted activities such as reading easier – this again leads back to the issue of implementation fidelity: characteristics of individual parents are likely to influence how they facilitate any intervention, and thus how effective it is.

The Abecedarian Approach. Initially developed in the USA in 1970s as part of a series of initiatives designed to improve social equality, the Abecedarian approach encompasses evidence-based teaching and learning strategies comprised of four interconnected elements: language priority, enriched caregiving, conversational reading and learning games, with the former element of prioritising language being a core component of all the others (Sparling & Meunier, 2019). The approach is designed to be facilitated at a group level by locally trained providers, yet delivered by parents to their own children. Adapted for Australian-specific use (the '3a approach,' University of Melbourne, n.d.) the strategies for prioritising language include talking face-to-face, imitating, using sophisticated vocabulary, using language contingently, asking questions and using pauses. All of these strategies can be incorporated into the other three elements, including enriched caregiving, whereby everyday routines can be harnessed as opportunities for enhanced responsiveness, including giving choices and talking about feelings. Conversational reading is an active and reciprocal form of reading involving pointing, questioning, and deviating from the script, and learning games involves 200 individualised, educational and 'game-like' interactions incorporating turn-taking, mark-making and matching shapes, amongst other things.

In Page et al. (2019), the researchers recruited 149 Australian children aged 24-56 months living across two remote Aboriginal communities (speaking thirteen different languages across the whole sample). Strategies were taught and facilitated via daily attendance at playgroup sessions which were ran by trained local providers. The researchers found that exposure to two elements (conversational reading and learning games) of the Abecedarian Approach Australia (the '3a approach') predicted stronger language development (here measured by Brigance Early Childhood Screen II), and that there were significant differences between dosage levels, with those children who received the highest dosage of the intervention receiving the most benefit. The researchers attributed the success of the 3a programme to the simplicity of its strategies (they did not rely on high levels of parental literacy in English), and the ownership and support provided by people in the local community, including the use of culturally appropriate adaptations (e.g., the translation of books into additional languages to support mothers). Whilst the strategies taught are not necessarily new or different to established good practice, the idea behind the Abecedarian approach was to bring these components together and practise them frequently and consistently, along with elements of individualisation, in order to bring about significant long term benefits for children (Sparling et al., 2021).

Effective Targeting and Tailoring of Parent-led Interventions

In Chapter 2, a multitude of factors associated with vocabulary outcomes were noted, including gender, maternal education (Dunatchik et al. 2018), co-existing social and emotional difficulties (Girard et al., 2016), and more proximal factors such as the number of trips to the library and parental attitudes to learning (Law et al., 2017b) - and many more. Some of these factors were tested in Phase One of the thesis against the LuCiD and MCS data, and the complex profile of results obtained supports the notion that accurately identifying those at risk of language delay is thus a complex task, as demographic risk factors, even well-established ones, do not guarantee adverse outcomes, and often intersect with other factors, as noted in Chapter 1. This raises the question as to whether large-scale and universally applied interventions, which have the potential to benefit all

children, would be preferable, rather than attempting to target a specific at-risk group. Greenwood et al. (2020) suggested just this, in that attempting to change the communication styles of vast numbers of individual parents and caregivers required population-level strategies similar to government campaigns employed for healthy living, such as tobacco health warnings and eating five a day.

In a report into language need and provision for pre-schoolers, Law et al. (2017a) suggested that targeted interventions were fairer in that universally provided interventions could unintentionally widen inequalities. Nonetheless, universal approaches still have a role to play, often in the form of quality-first teaching, or initiatives such as the National Literacy Trust's *Chat Play Read*, involving tips, ideas and videos on its website supporting parents with how they might do this, and why it is important. It is also worth noting that universal provision does not necessarily mean universal access (for example a family who reads regularly with their child may ignore a flyer containing a weblink to the *Chat Play Read* resources) – so a widening of inequalities is not a given conclusion. On the other hand, the opposite could be true, with families who already partake in language-enriching activities always looking to find new ways of doing so. Either way, universal interventions remove the problems associated with accurate identification of children in need of support, as noted above and by Law et al. (2017a), hence are worth considering as a viable option for parent-led interventions in some circumstances.

Where a the target group can be correctly identified, there is also the question of tailoring to consider, so that the likelihood of positive impact is maximised. In a systematic review of language interventions delivered to children from Latino or Spanish backgrounds in the US, Larson et al. (2020) found that the most effective interventions were those that had taken into account the relevant culture and/or language of the child, making either cultural or linguistic adaptations to the intervention. It is possible to apply these findings to a general approach for developing interventions, regardless of the target group in question - essentially that one size does not always fit all, and a tailored approach that carefully considers the culture of the group in question, across many different levels (local area, ethnicity, age of parents) may prove to be more engaging and more effective than those that do not, and could even go some way in alleviating the higher attrition rates seen within some communities (Walker et al., 2020). For example, tailoring an approach to an all-encompassing 'low SES group' is misinformed, as it does not consider the many individual subgroups with different cultural norms within such an umbrella term. Obviously, when an intervention is delivered universally to all children, with no sense of the target children involved, tailoring the intervention is impossible – however a range of materials and choices can still be provided. In addition, as noted above, interventions can be universal in that they are available to all, but not necessarily accessed by or beneficial to all, and so a universally available intervention could still have a target group in mind, and tailor the offering appropriately.

Summary and Limitations of Parent-led Interventions

A number of parent-led interventions have been reviewed, all of which focus on increasing interaction via reading, talking or playing strategies. Whilst interventions for this age-group could also take place in an early years setting, and one example was briefly considered (NELI), the focus here has been on parent-led interventions. This focus was due to parents being well-placed in a practical sense to spend 1:1 time with their children, as well as evidence which consistently supports the notion that

parental involvement is beneficial for children's developmental outcomes (Adamson et al., 2020; Law et al., 2017b; Wilder, 2013; Reece at al., 2010; Roulstone et al. 2012).

Overall, the studies reviewed support this suggestion, in that interventions which focused on improving parental responsiveness (specifically via the use of contingent talk), shared reading (specifically interactive approaches) or the use of visual prompting boards, all had a positive effect on children's language (and specifically vocabulary) and/or parenting behaviours which promote language learning. However, this support is not without caveat. Studies did not always measure children's language outcomes directly, and were sometimes focused upon parent behaviours (Kachergis et al., 2020). Some studies, notably McGillion et al. (2017) and de la Rie et al. (2020), involved a younger or older sample than the age group of interest for the current research, meaning that findings may not be generalisable to the current age-group of interest, as there may be different influences on word learning at play. Whilst some of studies (e.g., Noble et al. 2020; Burgoyne et al., 2018) predicted an effect of SES, in that their intervention would be less effective for those from a lower SES backgrounds, this was not found to be the case - although some studies found the opposite, in that the intervention was in fact more beneficial for this group (McGillion et al., 2017), and many had samples which were focused upon lower SES backgrounds (Feil et al., 2020; Burgoyne et al., 2018). Overall, this suggests that language-focused interventions have the potential to be beneficial for all children, and perhaps especially children from lower SES backgrounds. Finally, a number of studies (e.g., McGillion et al., 2017) noted the lack of long-term-maintained effects, and considered this might be due to the short time duration of the intervention being insufficient to change parental habits. Implementation fidelity also came up, in that it was often unclear whether parents had employed the strategies taught in a consistent way (Heidlage et al. 2018), and thus a broader issue arises in terms of what type of intervention (for example simple *versus* complex; structured daily routine *versus* ad hoc), would be effective in ensuring implementation fidelity and therefore increase the chances of positive outcomes for children.

In an argument for simplicity, parents vary hugely in their abilities, knowledge, parenting styles and attitudes towards learning, as well as their role in it. For example, Manz et al. (2010) suggested a mismatch between some parents' natural reading style, and the style which interactive reading demanded, and different attitudes towards reading and its function have also been suggested (Noble et al., 2020), thereby limiting the positive effect such an intervention could have. Despite being more readily available than other resources, parents' time may be limited, and they are influenced by a wide variety of demographic factors that will affect their ability to provide the optimal level of input. Such differences may provide a partial explanation for the variation in effectiveness of different parent-led interventions (Heidlage et al., 2018). The variability of parents suggests that interventions which are complex, effortful and require specific resources are more difficult for some parents to implement with fidelity. On the other hand, simple interventions that do not rely heavily on parental time, effort, knowledge or skill, may be more effective at instilling the long-term habits necessary to have long-lasting benefits on children's language outcomes (Sparling et al., 2021), with a final challenge involved in approaching the situation with care and diplomacy, and allowing for cultural adaptations where appropriate (as per the Abecedarian approach) to ensure parent buy-in. In summary, parent-led interventions have the potential to support children's language and specifically vocabulary development, but such interventions need to account for a wide range of parenting approaches, and should not under-estimate the difficulties in achieving a lasting shift in behaviour.

The Challenge of Changing Parental Habits: What Sort of Intervention would Work?

The preceding review has untangled a number of strengths and limitations of parent-led interventions. Key to addressing these limitations are the ideas of implementation fidelity and achieving long-lasting behaviour change which benefits children's language development. Facilitating such a change in parental behaviour and routine is not straight-forward, particularly when there might not appear to be an obvious or immediate gain for an individual, and imposing unfamiliar routines on families is unlikely to be effective (Mooney et al., 2016). Habits, defined as behaviours that are performed frequently in a particular context, and have become automatic over time (Wood & Rünger, 2016) are particularly important here. Even when particular habits are widely accepted to be linked with very specific negative outcomes (e.g., health risks), individuals may still choose to continue with such behaviours (especially if any potential consequences are a long way off) and retain a level of optimism bias (Sharot et al., 2011). If consequences are poorly understood, or not very tangible, as with children's language development, then this is even more likely. Parents and carers may not link their own habits (such as regular reading and contingent talk) to their child's future, and even if they know such things to be important, the repercussions may feel like they are far enough away not to worry about. Furthermore, bombarding people with negative information tends to be ineffective, and 'bad news' is often discounted (Moutsiana et al., 2013). In a Ted Talk based on her work as a Professor of cognitive neuroscience at University College London, Tali Sharot (TedX Talks, 2014) argues that we need to embrace this tendency for optimism by incorporating three principles into our strategizing: social incentives using positive language, immediate rewards and progress monitoring (e.g., informing taxpayers that 99% of people pay their taxes on time, or using a visual speed board as an immediate incentive to slow down). There appears to be some evidence of the former principle at work in the context of children's language interventions, that often use positive or neutral rhetoric (e.g., *Talk to You Baby* or *Chat Play Read*), although the use of social incentives, immediate rewards and progress monitoring is less apparent beyond the intrinsic reward of mastering the skills of oral language and reading.

The Nudge

Another approach that has been shown to be effective at inducing behaviour change in some contexts (and indeed could complement the general principles set forth above by Sharot; TedX Talks, 2014) is the 'nudge'. Although the concept has been around for a while, the term itself was first coined by Thaler and Sunstein in their 2008 book *Nudge: Improving Decisions about Health, Wealth, and Happiness*, and essentially endorses subtle adaptations to the environment which then influence or 'nudge' people's decisions – also known as 'choice architecture.' These adaptations should be easy and inexpensive to implement, and encourage people to make the more desirable choice by increasing the salience of such an option, or making it a default option, hence the descriptor: 'libertarian paternalism' (Gigerenzer, 2015, p.362). The principle of nudging is based on earlier dual-process theories of cognition and in particular the use of heuristics for decision-making (Tversky & Kahneman, 1973). Dual process theories posit two types of processing: automatic and nonautomatic, with the former heuristic-based and fast with no working memory resources required, and the latter deliberate, slow and effortful, and thus requiring resources from working memory (Evans & Stanovich, 2013; Melinkoff & Bargh, 2018). Traditionally the nudge was thought to capitalise on the automatic route, in that many decisions default to this mechanism unless people are willing to invest cognitive effort

(Evans & Stanovich, 2013) – however nudges have also been shown to be effective via the non-automatic route (van Gestel et al., 2021).

Nudges have been tested in a variety of contexts, with a frequently cited example of nudge theory in action being within the supermarket context: placing healthy foods such as fruit and vegetables near the checkout or at eye level, and unhealthy snacks in the furthest corners of the supermarket where they are less noticeable (Kroese et al., 2016). Individuals are still free to purchase these snacks, but some may be swayed by the extra effort involved, and others may simply fail to notice them in the first place. Another oft quoted example is portion size. In Venema et al. (2020) the nudge was in the form of a smaller sugar spoon causing tea-drinkers to add less sugar to their tea. Indeed, when the sugar portion size was reduced by 50%, the amount of sugar used by individuals reduced on average by 27%, although the effect was less pronounced with those deemed to have a stronger sugar habit. Logically, it appears stronger habits are harder to break with the nudge, whatever form that nudge takes (De Wijk et al., 2016).

Proponents of nudges argue that an individual's decision-making is not always considered and rational, therefore interventions to either increase the utility of the desired option or shift beliefs and attitudes will not always have the expected effect, and this may be particularly true concerning everyday contexts (Mertens et al., 2022). Mertens et al. (2022) reviewed 212 studies conducted over a ten-year period across a range of contexts and found an overall small to medium effect of nudge interventions, which was comparable to more traditional interventions such as educational campaigns or financial incentives. Nudges used in some contexts – for example food – were more effective than others – for example financial decisions. This coincides with the idea that everyday habitual behaviour may be more susceptible to nudges than bigger one-off decisions, such as buying cars and houses. The Mertens review was heavily criticised by Osman (2022), who claimed, amongst other things, that the publication bias was not accommodated for to the extent that it should have been. Indeed, Maier et al. (2022) found no evidence for the nudge being effective once the publication bias had been accounted for using a different statistical technique, although this criticism appears to be more directed at the utility of meta-analyses rather than the nudge itself.

Other criticisms of nudge theory operate on a more fundamental level, with some critics declaring that choice architecture manipulates and undermines people's autonomy in line with policy-makers' wishes (Ewert, 2020). What is optimal or desirable is often subjective, and questions arise as to who should make such decisions – but these criticisms could be applied to all types of intervention, not just the nudge. Sunstein (2022) concedes that from an ethical perspective, extra care may sometimes be needed to ensure vulnerable populations (e.g., children or those living in poverty) retain their decision-making power and are not unintentionally manipulated. Consideration should also be given to whether the nudge has benefitted the intended group of people, and whether the costs outweigh the benefits.

Relating Nudges to Improving Language Outcomes

The effectiveness of nudges as an intervention tool clearly needs further investigation, and yet even critics have conceded that nudges have the potential to work well when used in conjunction with other interventions, such as tax subsidies (Osman, 2022). In many contexts, including children's language development and parenting, statutory interventions via pricing or otherwise are neither

feasible nor desirable: offering financial rewards or making it legally enforceable for parents to provide an optimal home learning environment seem particularly unlikely options. This seems argument enough for pursuing a method that focuses on manipulating choice architecture (or 'nudging') – and in essence making the language optimal choice also the effortless option for parents. Such an approach could also incorporate Sharot's principles of focusing on the positive, and where possible, providing immediate reward and social incentives (TedX Talks, 2014). A good example of nudging within the context of children's language can be found in Ridge et al. (2015). Here, the researchers investigated whether the use of brightly-coloured signs displaying open ended questions (e.g., *Where does milk come from?*) placed around a supermarket would be effective at eliciting parent-child interaction. By discreetly observing 71 adult-child groups across three different supermarkets, they awarded an interaction score which included the number of conversational turns and the valence (positive, negative, neutral) of the interaction, finding that the presence of the signs resulted in significantly more conversational turns, most of which were of positive valence (although the signs were only found to be influential in supermarkets that had been deemed to serve lower SES areas). Whilst the authors agreed that short intensive parenting programmes could also be effective at promoting specific parenting techniques, which would in turn benefit children's language outcomes, they suggested that these were often costly, time consuming and subject to high attrition rates, hence novel ideas for intervention which harnessed the 'small everyday moments of communication could be one key way to support children's language development' (Ridge et al, 2015, p.129).

The National Literacy Trust's 'Literacy by Stealth' is a series of ideas based on a similar approach, in that the nudges take the form of incorporating language and literacy learning opportunities into activities and formats that children readily engage with anyway, such as video games or comics. Thus, by harnessing the power of children's interests, language and literacy are allowed in through the back door: by stealth, so to speak. This strategy is likely to be particularly effective with individuals who have struggled to engage with learning on more traditional platforms. The use of educational apps is also relevant here, both in terms of their contemporary technological quality having the ability to nudge and engage both children and parents, but also because of equality of access: 97% of families with children had access to the internet in 2022, either via smartphone technology or otherwise, with 20% of 3-year-olds owning their own phone (Ofcom, 2023) - whereas access to (particularly via ownership) physical books is still very much linked to income (Picton & Clark, 2023). A meta-analysis on the effectiveness of intervention studies using educational apps for children aged 3-9 years by Kim et al. (2021) found that the overall influence on literacy (and maths) was positive, although there was substantial variability between apps, and the effects were larger where the apps were developed for pre-school children, focused on the development of 'constrained' skills (e.g., letter knowledge), and evaluated via researcher developed outcomes rather than standardised measures.

Rowe et al. (2021) argued that the use of educational apps could 'transform' smartphones into opportunities for conversation and learning (Rowe et al., 2021, p.1). In this study, the researchers tested three interactive apps on 3 to 4-year-olds (n=76) to ascertain their influence on vocabulary, conversational turns and syntax, as well as the difference between providing the apps on their own, or with attached support and information for parents. All the parent and child linguistic measures (with receptive vocabulary being the chosen measure of vocabulary) increased significantly between the initial visit and three weeks later, regardless of which group they were in. All the apps tested in this study were developed with parent-child interactivity in mind, including the Animal Antics app,

which was designed to promote back-and-forth conversations by recording conversations and playing them back through an animation involving different animals and different settings. The potential effectiveness of apps for children's vocabulary and broader language skills is important, as it supports the notion that children, including pre-school children (Kim et al., 2021), can learn from less traditional platforms, as well as evidencing methods of intervention that work in harmony with parent routines. Whilst the apps tested by Rowe et al. (2021) were based around parent-child interaction, and not intended to be used independently by the child, there is still an element of nudging involved in that apps by their very nature are an easy and accessible format, even if there is still some effort required from parents.

Using Audiobooks to Improve Language Outcomes

One particular format that has the potential to nudge both children and parents towards an easy choice for learning is the audiobook. The benefits of parent-led interventions, particularly for birth to pre-school children (Adamson et al., 2020), were laid out earlier in this chapter, and they can be particularly beneficial and likely to facilitate longer-term benefits when the activity or message is easy to implement either through a clear structure and routine (e.g., Burgoyne et al., 2018) or a simple approach which does not rely on a huge effort or particular parental qualities, skills, or improvision (de la Rie et al., 2020; Sparling et al., 2021). The audiobook has the potential to fulfil these criteria, although the level of benefit will likely depend on to what extent they can facilitate parent-child interactions.

How will Listening to an Audiobook Improve a Child's Vocabulary? Constructivist and some hybrid models of word learning (e.g., Emergentist Coalition Model of word learning, Hollich et al., 2000) theorise that the quantity and quality of verbal input a child is exposed to is directly linked to the child's own subsequent vocabulary. Research evidence is supportive (e.g., Huttenlocher, 2010), which seems logical, as word labels cannot exist as innate knowledge, and children speak the native language(s) to which they are exposed. In a similar way to traditional shared reading, listening to the audiobook will expose children to vocabulary, and specifically to a greater diversity of verbs, identified as being important for overall lexical diversity by Phase One of this thesis. By listening to the audiobook, children increase the likelihood of acquiring the target verbs (and additional verb arguments, due to the nature of verbs tending to require additional surrounding words to make sense), and this likelihood increases the more they listen. As noted in Chapter 1, the more words children are exposed to, the quicker their processing speed becomes, inherently linked to their receptive vocabulary, but also leading to wider cognitive benefits and the development of all language skills, including expressive vocabulary (Fernald et al., 2006). As a child's vocabulary grows, new words become easier to learn as children are able to recognise familiar words more quickly (Weisleder & Fernald, 2013). Over and above the benefits of learning through exposure to words, audiobooks may also be able to facilitate parent-child interactions in the same way that other reading and play interventions have been shown to do earlier in this chapter, thus leading to even further advantages for children's language (Weizman & Snow, 2001; Fernald et al., 2006).

How Audiobooks may Appeal to Parents. Audiobooks draw parallels with traditional shared reading (e.g., by increasing a child's exposure to vocabulary, and possibly facilitating parent-child interactions), and thus have the potential to share some of the well-established benefits associated with such a practice, for example reading for enjoyment (Pillinger & Wood, 2014), and exposure to diverse and

rare vocabulary (Beals, 1993). Furthermore, some of audio's characteristics may provide benefits which go beyond sharing physical reading books, and in doing so address some of the limitations of parent-led interventions already discussed: namely, that they require minimal effort, resources and skill, and their simplicity may increase the likelihood of them transcending the duration of the intervention programme. One important quality, particularly related to the time resources a parent has available, is the portability of audiobooks, which, in the age of the smartphone, can be played wherever and whenever a busy schedule allows for it - in the car, bath and whilst multi-tasking (Picton, 2018, as cited in Best, 2020). With respect to knowledge and skill, Woolfson (2008) commented that audiobooks are particularly suited to classroom contexts with a range of abilities, as they allowed all students to access the same text, even if this was beyond their fluent text reading ability. The same argument can be applied to parents who lack the confidence to read with their children (and who may be struggling readers themselves) – audiobooks do not depend on their reading ability, and provide access to a wider range of books (Miller, 2010). Parents and children are likely to feel the same sense of achievement as when finishing a traditionally formatted book, something that can be aligned with Sharot's principle of immediate reward (TedX Talks, 2014).

Data quoted by the National Literacy Trust suggests that many parents do not read with their children on a regular basis, with 45% of children aged 0-2 years being read to daily or almost every day, with this figure rising to 58% for 3 to 4-year-olds, before dropping down to 44% for the 5-7 years age group (David, 2020). Whilst this may be due to the reasons discussed above such as a lack of knowledge, skills, or time, it might also be because of a lack of motivation and interest in reading, and thus a reluctance to put in the effort involved. Audiobooks also address this point, in that they can be accessed easily with minimal effort required (although the limitations of listening without interaction are discussed below). Not only that, but audio may present itself as a preferred format for some parents, thus creating a renewed interest and enjoyment in books which is likely to have positive repercussions for their child. This was found to be the case in Zientarksi and Pottorff (1994), who completed an action research project in a secondary school with a group of struggling readers and found that they reacted enthusiastically and with renewed interest in literacy in post-listening discussions. Demographically, audiobooks may also be particularly appealing to younger parents who have grown up with increased engagement with technology (Waite, 2018), although consumption of audiobooks (as with physical books, see Clark & Picton, 2018) is still higher amongst higher income brackets than lower ones (Best, 2020).

Audiobooks may be particularly relevant to families from underrepresented groups, and to those experiencing multiple pressures or from lower SES backgrounds, where previous research (e.g., Mol et al., 2008) has shown that more traditional interactive reading strategy interventions may be less effective (although this was not supported by the results in Burgoyne et al., 2018, which suggested that interactive reading strategies could be effective across a range of SES backgrounds). Nonetheless, for families experiencing multiple pressures, be they financial, time-related or otherwise, by offering a language intervention in the form of an audiobook that is both minimal cost and effort, parents may be nudged into supporting their children's language development in an effortless way. In certain circumstances, such an intervention could be the difference between children receiving some level of quality verbal input, or none at all. Similarly, with respect to underrepresented groups, as the audio holds a unique position in its ease of adaptation, it presents an ideal way to relate to parents and children by amending aspects of the story and its delivery to make them more relatable for specific

groups. Self-Determination Theory (Ryan & Deci, 2000) proposes that humans have three fundamental needs that need to be fulfilled to achieve personal growth: competence, autonomy and relatedness. All three of these components can potentially be achieved via the audio intervention and thus motivate parents towards the activity – relatedness via ensuring content is relevant to participants' lives as described above, competence for parents who are uncertain about their own literacy levels, and autonomy via the independence it confers on the child being able to 'read' an audiobook without adult assistance (again, whilst acknowledging the limitations of any benefits which can be achieved where there is no parent-child interaction at all).

Evidence for Audiobooks Improving Children's Language Outcomes. Wolfson (2008) suggested that audiobooks were particularly helpful for adolescents, including those who had additional needs, were reluctant readers or who had fallen behind. Audio developed listening and pronunciation skills, as well as facilitated best practise modelling of reading in general, and was able to 'support the development of all four language systems: phonological, semantic, syntactic and pragmatic' (Wolfson, 2008, p.106). Rogowsky et al. (2016) tested the influence of learning modality on the verbal comprehension of 121 adults, finding no significant difference between the three groups (reading, listening, reading and listening together), with results holding between males and females, and at both time points (immediately after the information was presented, and two weeks later). These results suggest that different modalities of learning, including audio, produce similar outcomes for learning - although the results pertaining to adults with well-developed listening and reading comprehension skills are difficult to generalise to early stages of children's language acquisition, as acknowledged by the researchers, and further research is thus required.

In Olson et al. (2024), 314 US-based 3rd and 4th graders (aged 8-9 years) participated in a RCT with three conditions (audiobook only, audiobook plus scaffolding in the form of facilitators checking progress and teaching some vocabulary items and comprehension strategies, and control) aiming to explore the effectiveness of an audiobook intervention on vocabulary learning, and ultimately how this vocabulary then influenced their reading comprehension. After 8 weeks, significant improvements were found between pre- and post-test receptive and expressive vocabulary for both audiobook conditions but not the control group, but only when proximal, audiobook-specific measures of vocabulary were concerned, as opposed to standardised measures. The proximal measures were designed to measure children's knowledge of the words which appeared in the audiobooks, so were more likely to capture changes than the standardised measures, however the researchers suggested that a correlation between the two measures at pre-test indicated that meaningful variation in children's vocabulary was still apparent, and future research which was over a longer time period had the potential to generate larger transfer effects and thus more tangible effects on standardised measures. Disparities were also apparent between differing SES backgrounds (operationalised by parental education) and reading abilities; good readers benefited significantly from both audiobook conditions, whereas less proficient readers, and those from a high SES background, only benefited from the audiobook if it was accompanied by scaffolded support. Unexpectedly, the opposite was found for those from low SES backgrounds, in that their vocabulary benefited in the audio only condition, but not when scaffolded support was provided. The researchers offered the context of data collection during COVID-19 as an explanation for this result, in that lockdowns may have resulted in less access to quiet spaces in the homes of children from lower SES backgrounds, hence making the scaffolded support less effective. On the other hand, a general lack of access to reading resources

could also have been a reason why the provision of the audiobook on its own was enough intervention to benefit this group's vocabulary.

Whilst the children in Olson et al's study were considerably older (mean age: 9.47 years) than the children of interest in the current research, the study nonetheless clearly supports the use of audiobooks for vocabulary learning. Of particular interest, whilst unexpected, was the result that some children benefited from the audiobook-only intervention, albeit with on-screen text as standard in both groups (thus the relative influence of the visual and audio input is unclear). For the current study, the finding that children's vocabulary may benefit from audio alone, as an independent activity, without any scaffolded support, is promising. This is not to suggest that this is the optimal way to access audio material – indeed the overall recommendations of Olson et al. (2024) are for a combination of text-supplemented audiobooks *with* personalised support, particularly for struggling readers – however where this support may not be available, children's vocabulary can still benefit from audiobooks. Audiobooks also present an ideal format for introducing new vocabulary, and thus for the purposes of the current research, resources could be easily adapted to include the all-important target verbs whilst still retaining the normal sentence (as opposed to nonsense) structure. The audio format is also ideal for adapting content or other aspects of delivery with each cycle of feedback.

Some Limitations Associated with the Use of Audiobooks. The use of audiobooks, and specifically the use of audiobooks to introduce new vocabulary, is not without its limitations. Listening continuously lacks the interactivity of reading (although this is not necessarily the case, as traditional reading can also be done without much interaction), and if listening to audiobooks replaced traditional reading, then this would deskill older readers at the decoding process, and give younger readers less opportunity for developing print awareness. Without these elements of reading, and the presence of a physical book, listeners may find themselves less likely to pay attention – especially if listening in a distracting environment or to an unfamiliar dialect (Waite, 2018). A physical book provides more autonomy for an individual, in that they control when they turn the pages (Varley, 2022, as cited in Best, 2020), how long they look at a picture for, and what the 'voices' sound like. An audiobook also lacks the experiential component of learning in that it lends itself to a more passive style of learning rather than an active 'learn through doing' approach which is advocated by numerous developmental theorists (e.g., Kolb's 1984 experiential learning cycle) and educational approaches (e.g., learning through play in the EYFS; Pie Corbett's talk for writing scheme). The importance of experiential learning was further evidenced in Duncan et al. (2019), where a movement and storytelling intervention was tested via a cluster-randomised control trial involving nine schools and 214 children, with findings indicating small but significant effects on language, particularly expressive language. For audiobooks, this inherent lack of experiential qualities is a limitation, but can to some extent be mitigated by actions which take place alongside the audiobook, facilitated by physical props, parents or otherwise.

Whilst the noted lack of interactivity inherent in the audiobook format could be perceived as a strength (in that it may appeal to parents who face barriers to interacting), it is important to note that all of the parent-led interventions reviewed earlier in this chapter were based on the premise of increasing and/or improving parent-child interactions, and how such interactions benefit children's language. Whilst the audiobook may result in parent-child interactions, as noted above, it can also be

listened to alone and without interaction, and even where parent-child interaction ensues, there may be less opportunity for contingent talk, noted as important for children's language development (e.g., McGillion et al.,2017). Due to the focus of young children's attention often being held by physical objects, contingent talk lends itself to situations where physical entities are involved, for example playing with toys or reading with a physical book where words and pictures can be referred to. Listening to audiobooks is not physical in nature, so whilst they are still capable of holding the attention of a child, and thus could elicit contingent talk if an adult were present, this is more difficult without physical referents both in real time and in the real world (Tamis-LeMonda et al., 2014). This could perhaps be somewhat mitigated if an audiobook were to be provided with an accompanying text or character toy, in that such an item could be the object of a child's attention and form the basis of any contingent conversations.

Ideally, audiobooks should supplement existing practices of traditional reading and other quality verbal interactions, rather than replace them (Best, 2020), as optimal language learning is thought to be fuelled via the interactive nature of contingent back-and-forth conversation (Goldin-Meadow et al., 2014; Tamis-Lemonda et al., 2014), as mentioned in Chapter 1 when discussing the different qualities of child-directed speech (CDS) and overheard speech. On the other hand, the limitations associated with audiobooks do not suggest that children cannot learn effectively through this format at all (see Olson et al., 2024, above), only that it might be a less effective method than learning experiences with more interactivity. Whilst the audiobook format does not necessarily lend itself to back-and-forth conversation, neither does it fall squarely into the category of overheard speech either, which is argued to be more difficult for young children to learn from (Golinkoff et al. (2019). Indeed, some of the characteristics of CDS such as simplicity of content and exaggerated prosody should be equally present within an audiobook as they are in real-time CDS. The use of accompanying physical items, as mentioned above, could also help facilitate any subsequent interactions with caregivers which happened outside of listening time, developing the child's use and understanding of decontextualised speech (also important for language learning, see Rowe, 2012). Physical accompaniments could also render the audiobook itself more contingent in nature, in that there would be a referential triangle of sorts between the child, the related physical item (e.g., a toy character from the audiobook), and the audiobook itself, all focused on the same content.

Summary of Chapter 7 and Key Issues to be Addressed by Phase Two

This chapter began by reiterating the context of this thesis, and the importance of addressing the vocabulary gap (how interventions are a way of doing this), and anchoring Phase Two to the outcomes of Phase One. Interventions were defined and their prevalence considered in general, as were important themes such as intervention timing and how other factors may play a role in whether an intervention is effective or not, over and above its design and content. Whilst interventions within early years settings were briefly considered, the crucial role which parents play in their child's development (particularly in the pre-school years), including their cost-effectiveness, availability, and a motivation for their children's well-being, were offered as a justification for focusing on reviewing parent-led intervention studies specifically.

Overall, evidence suggests that parent-led interventions can be effective at supporting children's language. The interventions reviewed, whether primarily focused on reading, playing or responding, all had an element of parent-child interaction, and this was also true for the selection of

educational apps considered. Despite the generally positive effects on language and specifically vocabulary of parent-led interventions, longer-term effects on language were not always maintained (McGillion et al., 2017), or whether they were or not was unclear due to a lack of follow-up testing (Heidlage et al., 2018). This latter point seems particularly crucial. If an intervention is going to be successful, then it needs to facilitate wide-ranging attitude and behaviour change that extrapolates to other situations and continues beyond the short-term benefits that may be apparent after a short and focused intervention. Habits, defined as frequently performed behaviours which have become automatic in certain contexts (Wood & Rünger, 2016) are particularly important here, in that an intervention that manages to facilitate the formation of new habits would be particularly valuable. The short duration of intervention programmes was discussed as a possible explanation for the potential absence of long-term benefits for language interventions, as were difficulties in ensuring implementation fidelity amongst parents. Short programme lengths could be less capable of shifting parental habits in the long term, and with respect to implementation fidelity, there was no guarantee that parents were consistently sticking to the strategies taught by the intervention, and even when they did, wide variations in parental skills, knowledge and characteristics meant that differing results may ensue. An intervention which was simple, and did not rely overly on parental skills, knowledge and characteristics, may be more likely to shift habits in the long term.

In light of some the limitations of parent-led interventions discussed in terms of achieving long-term change and benefits for children's language, habits were discussed in the context of the 'nudge,' whereby it is argued that subtle adaptations to the environment are capable of nudging people into making 'desirable' decisions, an idea linked to earlier dual process theories of decision-making. In pursuance of this approach, similar in style to the NLT's *Literacy by Stealth* (whereby language and literacy learning opportunities are incorporated into children's existing routines and interests, and through varied formats: 'through the back door'), educational apps and, ultimately, audiobooks were considered as potential formats for an intervention. Evidence supporting the use of audiobooks was reviewed, and how they might appeal to parents and overcome some of the limitations of parent-led interventions discussed. For example, audiobooks have many of the benefits of traditional shared reading (e.g., exposure to a diverse range of language), but rely less on parental time and skill, and the alternative format may appeal to those who have struggled to establish regular shared reading routines. Associated limitations were considered, notably the lack of experiential quality and opportunities for contingent talk and parent-child interaction in general – as well as how these might be mitigated.

The main aim of this Action Research study is therefore to develop a suitable audio resource that can be utilised as a parent-led language intervention, and which has the potential to circumvent some of the challenges associated with parent-led interventions discussed, and therefore allow for and longer-term changes in habit and resulting benefits for children. Despite the scarcity of existing research evidence, which is often focused on older age groups - the use of audiobooks is increasing, both inside and outside the classroom (Eyre, 2018), and it is a format which has the potential to be harnessed. As discussed above, audiobooks have the ability to appeal to parents and improve children's language outcomes simultaneously, with the already discussed caveat that such improvements may depend to what extent the audio facilitates parent-child interaction. Nonetheless, this thesis is set against the context of only 58% of 3 to 4-year-olds regularly taking part in shared reading (according to data quoted by the National Literacy Trust: David, 2020), so it is also worth

considering the possibility that audio may provide some benefit even in the absence of parent-child interaction. Ultimately, the idea that audiobooks may have the potential to nudge the parent and child into ‘good’ language habits that endure beyond the length of the intervention, allowing the child to expand their vocabulary range and diminish any so-called 'vocabulary gap,' will be explored by Phase Two.

Chapter 8: Method (Phase Two)

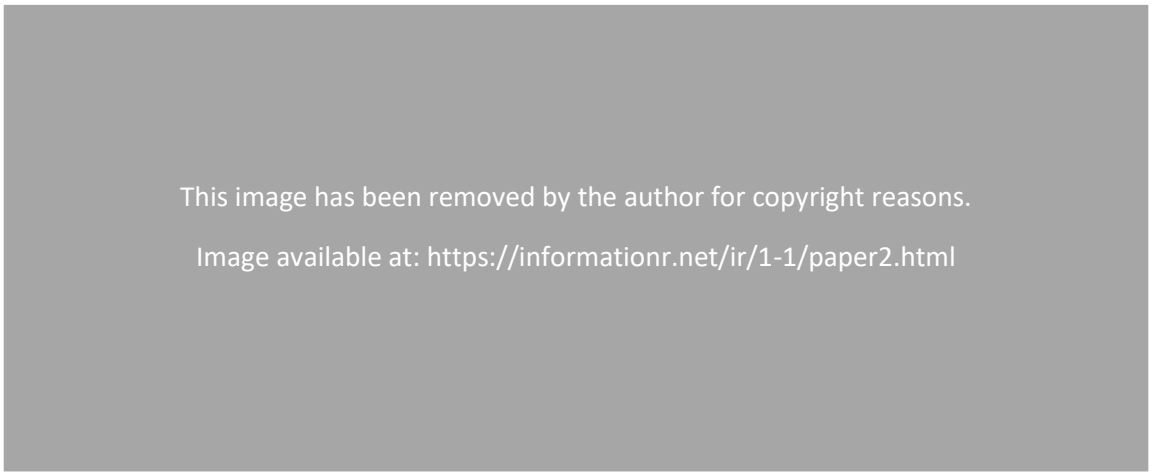
The starting point for this small-scale Action Research project is the culmination of the outcome from Phase One of this thesis (highlighting the importance of verbs for pre-schoolers' vocabulary development), and the ongoing challenge of ensuring all children are exposed to the vocabulary input that they need in order to develop and expand their own vocabularies, readying them for school and beyond. In consideration of the literature reviewed on parent-led interventions in the previous chapter, and the limitations associated with such interventions, the initial idea is therefore to provide participant parents and their children with an audio intervention designed to help build children's vocabulary, and use Action Research methodology to facilitate the incorporation of parents' opinions in order to help develop this intervention for future use. Using templates from the Early Intervention Foundation, a theory of change, logic model and intervention blueprint were completed to help further rationalise why such an intervention might be needed, and how resources and activities would lead to specific outcomes in the short, medium and longer-term (see Appendices I-K). The research question is thus: how can a vocabulary intervention be effectively incorporated into an audio format? At this stage of developing the idea, and to maximise participation, the intended participants can be any parent of a 3-year-old child. The remaining part of this chapter will define Action Research (AR), before considering a number of quality criteria as a framework for formulating and evaluating the current project. Within this framework a range of key methodological issues will be considered, including epistemological stance, suitability and limitations of using AR for the current project, details of sample and resources, methods of data collection and analysis, and construction of the written report.

What is Action Research?

Action Research (AR), as the name suggests, is about doing things (action) and generating knowledge (research) which is of practical importance. It is a flexible and usually qualitative methodology concerned with 'worthwhile human purposes' (Reason & Bradbury, 2001 p.1), and is based upon iterative cycles of planning, acting, observing and reflecting (see Figure 1). AR acknowledges the central, non-objective role of both the researcher and the participants, who should be considered as co-researchers rather than passive subjects (Bradbury, 2015). Indeed, subjective experiences of both researchers and co-researchers in terms of their values, intuitions and relationships are all valued and included as part of the process (Boulton & Preiser, 2021).

The cyclical nature of AR described above and depicted in Figure 1 tends to be common across all variations of AR, however it is important to note that researchers have conceptualised AR in a variety of different ways, and there is no single agreed format for how AR should be done, who it should be done by, and how it should be presented (McNiff & Whitehead, 2005). One particular disparity is whether AR should be a purely practice-based approach carried out by the professional who wishes to improve their practice (known as first-person AR), often in a teaching context (McNiff & Whitehead, 2005), or whether AR can be carried out by an external researcher (Greenwood & Levin, 2006). The term 'second-person practice action research' is appropriate for the current AR project, which refers not to improving one's own practice per se (as with first-person AR), but 'engagement in collaborative work in co-inquiry and shared action with others on issues of mutual concern, through face-to-face dialogue, conversation and joint decisions and action' (Coghlan & Shani, 2021, p.468). The

notion of shared action and joint decisions requires finding common ground between researcher and co-researchers, and an there is an emphasis on shared power and equality in contribution from all inquirers (McArdle, 2004).



Quality Criteria for Action Research Projects

Reliability, validity, replicability, objectivity and generalisability are common terms used to evaluate the effectiveness of quantitative research, but are often considered inappropriate for evaluating qualitative research, including AR projects (McAteer, 2013). In contrast, a variety of alternative quality criteria have been suggested for AR. Reflecting on these criteria is an important consideration, not just as a means of evaluating the quality of the project retrospectively, but also to use as a guide before and during the research process. Prior to delving further into a specific set of quality criteria suggested by Feldman (2007), it is first necessary to discuss questions of epistemology.

What can be Considered Knowledge?

Epistemology is concerned with knowledge, and understanding what it is to know (Stanford Encylopedia of Philosophy, n.d). It is an important consideration because knowledge forms the basis of any claims which are made based on research findings. What can be considered knowledge in the context of AR is discussed by Feldman (2007), who refers to the opposite extremes of naïve views of realism, where claims of truth or knowledge must be agreed upon and true beyond all possible doubt, and radical constructivist perspectives, whereby different truths exist in the mind of each person, and therefore there is no way of knowing what is real, as we cannot step outside of this to verify any knowledge or truth (Von Glaserfield, 1993). Feldman (2007) argues for a middle ground constructivist position, whereby some claims to knowledge can be accepted on face value, based on persuasiveness and credibility, whereas others need more evidence and consensus around whether the claim in question does in fact constitute knowledge and truth. In this journey towards knowledge, Coghlan and Shani (2021) suggested that the process of abductive reasoning was key, whereby conjecture and reflection in the present were able to produce 'tentative answers' and 'exploratory hypotheses' (p.467). Finally, Winter (2002) argued for the importance of addressing authenticity, in that it was inextricably wound up with knowledge and truth. Authenticity, and presenting the genuine voice, could be achieved via the principles of reflexivity (a self-questioning approach) and dialectics (a

collaborative approach recognising multiple perspectives), both of which will be considered in more detail below.

Quality Criteria suggested by Feldman (2007)

As previously noted, a variety of quality criteria for AR have been suggested (e.g., Oancea & Furlong, 2007; Heikkinen et al., 2007), and the epistemological concerns discussed above, regarding what can be considered as truth and knowledge, feed directly into these criteria, in that the knowledge sought to be generated by AR must seek to be true and credible. These criteria are not necessarily mutually exclusive, and can be considered in unison (McAteer, 2013). Indeed, the criteria set forth by Feldman (2007) embrace this collaborative approach by proposing adding the principles of reflexivity and dialectics suggested by Winter (2002), and built on by Heikkinen et al. (2007), to four further interwoven indicators of quality. In contrast to other writers on the topic (e.g., McAteer, 2013), Feldman (2007), argued that the principle of validity was in fact relevant to qualitative research, including AR. Quoting Hammersley (1992, p.69), he suggested that the concept of validity was broader than simply how well something has been measured, and extended to accurate representation via describing, explaining or theorising – all of which could be done via quantitative or qualitative methods. Such a broad interpretation of validity required the 'middle ground between naïve realism and radical constructivism' to be taken, as described above (Feldman, 2007, p.24). The quality criteria suggested by Feldman (2007) will now be considered in turn, both in general, and in respect to how they were addressed by this particular AR project.

Reflexivity Principle

Reflexivity is the reflective component of AR, and how such reflection and self-questioning (Winter, 2002) enable the cyclical nature of reviewing previous actions and planning subsequent ones to progress. Beyond self-questioning, Heikkinen et al. (2007) proposed that reflexivity should also require the researcher to examine their relationship with the research aims, question their presumptions of knowledge and reality, and make materials and methods transparent. Furthermore, Feldman (2007) adds that some attempt should be made to explain why the account given in a particular AR report is more true than any other, as well as how the principle of reflexivity has been operationalised.

For the current AR study, questions of knowledge and reality have been questioned under the subheading *What can be considered knowledge?* with the view taken similar to that of Feldman (2007) and aligning with the overarching pragmatic approach of both phases of this thesis: the level of evidence required to support a claim of knowledge depends on the claim being made, with some requiring more than others to achieve the authenticity strived for by Winter (2002). The reflexivity principle was also operationalised via conversations with a critical friend in the form of the researcher's PhD supervisor, with further validation given by the wider supervision team. A critical friend in the context of AR is someone whose opinion is valued, and who's role is to listen, support and challenge (McNiff, 2002). A validation group provides additional critical feedback on work done, and whilst not falling within McNiff's definition of 4-10 people (McNiff, 2002), the researcher's wider supervision team nonetheless provided the suggested critical feedback at various stages of the project. Heikkinen et al. (2007) also suggest that it is important for the researcher to examine their own relationship with the objects of the research being done, as this will necessarily be shaped by the

researcher's own experiences and identity. This aspect of reflexivity was contemplated below within the consideration of the researcher's position as a stakeholder.

Dialectic Principle

The dialectic principle is concerned with the inclusion of multiple perspectives, and together with reflexivity, can add to the authenticity of AR (Winter, 2002), with authenticity already noted above as being interwoven with issues of truth and knowledge. Feldman (2007) proposes that the dialectic principle incorporates more than just including different voices and perspectives, but was a 'critical exchange of ideas, claims and counter-claims (Feldman, 2007, p. 28). From an epistemological stance, the creation of knowledge via multiple perspectives fits in with Feldman's middle ground approach, in that it accepts that we need some 'level of correspondence [with others] with an agreed upon reality' (Feldman, 2007, p.24) to go about our daily lives, therefore supporting the use of collaboration and multiple perspectives.

Beyond epistemological considerations, it is important to consider how the dialectic principle, as described, can be demonstrated in practice. McArdle (2004) referred to an emphasis on shared power and contribution, and the finding of common ground between researcher and co-researchers – but how can this balance be achieved between different voices and perspectives? Achieving this balance is particularly pertinent between the researcher and the participants/co-researchers, for example when considering who is responsible for the actions and who for the generation of knowledge. Ideally, both responsibilities are a collaborative process, and yet it is important to acknowledge that in practice this will be an imbalance: the researcher has designed and implemented the study, and is solely responsible for the writing up, which could lead to a biased representation of what has happened, and of the actionable knowledge which has been created.

In an attempt to circumvent the risk of bias and imbalance described, and to operationalise the dialectic principle, this AR project made a number of commitments. Firstly, all the data obtained was presented, as suggested by McAteer (2013), even if this felt at odds with the researcher's perspective, or was in contravention of the majority held view (also known as negative case analysis). To facilitate this, a broad range of data was gathered by semi-structured interviews, with the aim of mirroring normal everyday conversations. In the pursuance of this conversational aim, no recording took place (as in Dom & Willockx, 2022). The initial narrative of the change process was presented in a linear fashion to ensure each participant had a broadly equal weighting with respect to the data presented. Secondly, to ensure the 'critical exchange of ideas' proposed by Feldman (2007, p.28), and to achieve an equal balance of perspectives, each idea presented was reflected upon both independently and via utility of the critical friend relationship. Finally, this AR report held itself accountable to the abuse of narrative privilege. As previously noted, AR methodology embraces the subjectivity of the researcher, and data such as diary entries and personal notes are legitimate data, and these, alongside personal reflective comments, can and should form part of the narrative. On the other hand, it is important that this narrative privilege is not misused (Adams, 2008). One way of ensuring this is to embrace the principle of dialectics and incorporate different perspectives of the same event, thereby ensuring that the overall narrative is as reliable and trustworthy as possible, and a true representation of the situation (Bold, 2012). Care should also be taken to strike a balance between embracing subjective opinions and preaching the moral high ground. This is particularly pertinent when dealing with information deemed to be sensitive or likely to be judged by others.

Whilst the focus of a particular AR project might be an intended topic, in practice many other aspects of participants' behaviour may 'come under the microscope' (Denscombe, 2017, p.128). For the current AR project, the intended focus of the interview was the content and format of the audiobook, however it was conceivable that other aspects of co-researchers'/participants' lives might have been discussed in relation to this, such as reading habits and parenting in general, which could have veered into the sensitive category. Due care was taken when presenting such information.

Methods of Resource Construction, Data Collection and Analysis

In his third quality requirement, Feldman (2007) suggests that a clear and detailed description is needed for how and why the data has been collected. With respect to the latter requirement, detailed reasons have been laid out in the first paragraph of this chapter, and to summarise, parent data collected over the duration of this AR project was for the purpose of facilitating the development of a simple audiobook resource intended to improve children's vocabulary. The audiobook format was chosen after a review of the literature of parent-led language interventions in Chapter 7 supported the use of parent-led interventions, but suggested limited evidence for longer-term effects, and variation in the effectiveness of interventions, particularly when fidelity requirements were high. The audio format presented itself as a potential solution, although limited evidence regarding the effectiveness of audiobooks, particularly in younger age-groups, required further investigation. How the data was collected will now be considered, including descriptions of how the resource was created, the sample, and the process of data collection and analysis.

The Audiobook Resource. The audio resource was created by adapting a freely available and copyright free script of the classic children's fairy tale *Goldilocks and the Three Bears*. Initially, the script was only adapted to include the target verbs, but in later cycles further adaptations were made after reflecting upon parental feedback and literature. Despite the aim of this AR investigation being resource development (rather than evaluation), a decision was made to create two versions of the audio resource to reflect how a control and intervention group would look in a later efficacy trial. The 'easy' version contained a narrower selection of repeated verbs (e.g., *was, came, said*) and therefore had a lower verb diversity. The 'harder' version contained a more challenging selection of non-repeated verbs (e.g., *gobbled, stretched, climbed*) and therefore had a greater verb diversity (see Appendices L and M for initial versions of audio scripts). In line with the Phase One of this thesis, 'verbs' are considered to be any word that parents of young children and adults working in early years settings understand a verb to be: an action or 'doing' word. This approach allows different forms of the same verb to be counted as different verbs (e.g., *eating* and *ate*). Each version was approximately 3 minutes long and contained a very similar number of word tokens and verb tokens in total (see Table 1). The participants were assigned to the 'easy' or 'hard' version in order of their recruitment.

Table 9

Summary of Number of Types and Tokens (Total and Verbs only) for each Version of the Audio Script

Version	Total words (tokens)	Total verbs (tokens)	Total verbs (types: lemmas)	Total verbs (types: different forms)
Easy	458	67	22	28
Hard	462	69	43	46

Once the scripts had been finalised, they were then recorded using an adult male voice with no marked regional accent, and transferred to a sound file that could be sent directly to the participant's phone via email, with the intention that it could be played from the phone (or indeed any other compatible device that could access the email). One of the cycles involved introducing story-related finger puppets (see Appendix N), and the final recording used a different narrator, but no further resources were used.

The Sample. Seven parent-child dyads were recruited via opportunity sampling (specifically via the local pre-school and the researcher's contacts). All children were 3 years old at the time of recruitment, with no other constraints placed upon who could take part. To fit in with the cycles of AR methodology, recruitment and participation in the research took place in stages, with initially only two parent-child dyads being sent the audio files (1 easy, 1 hard), feedback sought, and action taken. The actions taken took the form of amendments to the scripts and audio files made, or additional resources being provided, although there was no prior stipulation of what form the action needed to take. Once action had been taken, the next two parent-child dyads were then sent the amended versions of the audio files, plus any additional resources if applicable. Any changes made were done so to both versions of the audio resource.

No participant information was collected beyond names (pseudonyms are used here) and age of child (3 years old, to allow for consistency with Phase One of the thesis). No further participant information was required due to the aim of this Action Research being exploratory; all comments and suggestions made by parents were of interest, without recourse to related demographic information. This approach was in line with GDPR principle of data minimisation, whereby the only data that should be collected is that which is directly relevant to the aims of the research. Despite Phase One of this thesis linking specific demographic factors to children's vocabulary diversity, a decision was made to exclude this information from the current research design. The reasons for this are that first, results from the demographic variable analysis in Phase One were inconclusive and did not triangulate successfully between the different datasets used, and second, because at this exploratory stage, it was not desirable to limit the sample of potential participants. This is not to say that future research will not focus on targeted groups – indeed it is highly likely that it should in order to provide a more targeted and tailored approach.

Data Collection. After initial recruitment, parents attended a brief meeting in which the aims of the study were explained and there was an opportunity for questions. During this meeting parents were instructed to play the audio to their child as many times as they liked, but a minimum of once per day, during a 1-week period, noting down any thoughts they had. These thoughts could be anything at all, but comments about specific features of the audiobook, the child's engagement and ease of delivery were all suggested as examples. At the end of the one-week period, a conversation was then had with the participant parent to discuss their feedback, which was then fed into the next cycle of change. Meetings took place in person or otherwise to suit the needs of the participant, and lasted approximately 30-45 minutes. Interviews were not recorded, but notes were taken and extracts from these notes are provided throughout Chapter 9 (findings and discussion).

The Data. As noted above, data for this AR project consisted of notes taken by the researcher during conversations between the researcher and the parent participants/co-researchers. The majority of these notes constituted an abbreviated version of the parent's words, but also included reflective comments from the researcher. 'Data' within an Action Research project can encompass all items of data collected, with evidence defined as the argument and claims built around a particular set of data (McAteer, 2013, p.63). For AR, Dadds and Hart (2001) assert that methodological inventiveness can and should be employed in this respect, and in support of this assertion, data can exist in numerous forms: anecdotes, diary entries, photographs, videos, children's drawings, notes or transcripts of interviews or discussions – and much more – are all acceptable formats (McAteer, 2013; Feldman, 2007). Where data is from interviews or discussions, as for the current AR project, recordings, full transcripts, partial transcripts or note-taking are all acceptable ways of formatting the data (e.g., Einboden et al. 2023; Dom & Willockx, 2022), and the prevailing consideration is how well immersion within the data can take place, rather than its format (McAteer, 2013). Data was collected via note-taking (as opposed to recording) with a view to mirroring an everyday conversation, as well as avoiding possible inhibitions within that conversation, for example potential parental sensitivities when discussing their children's development in the knowledge that such conversations were being recorded.

Data Analysis. Data collected in AR projects is usually (although not always) qualitative in nature, and thus the methods of data analysis tend to be concerned with identifying patterns and themes within the data. For example, Xin and Brion-Meisels (2022) used a combination of first etic and then emic coding to analyse interview transcripts, whereby the former involved using themes from the literature to identify chunks of relevant data, and the latter allowed for the inclusion of themes which were not previously predicted, and had emerged from the data. Another ubiquitous method within the context of AR is Thematic Analysis (e.g., Ramlal & Augustin 2019; Chen et al., 2021; Acharya et al., 2020). Thematic analysis can be either inductive (allowing codes and themes to emerge from the data) or deductive (when the codes and themes have been pre-decided, based on existing theory). For the current AR project, inductive thematic analysis (TA) was utilised; a flexible and exploratory approach whose primary aim is to uncover meaning (McAteer, 2013). Inductive TA also aligns with the earlier quality indicators discussed, such as the recognition of multiple perspectives (dialectic principle), being open to new ideas via a collaborative and reflective approach (reflexivity principle), and the 'middle ground' notion of what constitutes knowledge as put forward by Feldman (2007), whereby claims of knowledge can be both accepted at face value, and necessitate further evidence and agreement – the latter of which can be facilitated via the searching for themes across participants as required by TA. The flexibility of TA means that it can be applied to most types of data, including field notes (Chen et al., 2021) photographs, videos, and other visual material (Braun & Clarke, 2019), as well as allowing for some flexibility in the writing up process, which will be discussed in more detail below. Whilst the process is inherently iterative, the following steps laid out by Braun and Clarke (2006, p.87) provide a framework for analysis:

1. familiarisation with data;
2. generate initial codes;
3. search for themes;
4. review themes;
5. define and name themes;

6. write up.

Coding procedures, whilst both semantic (surface) and latent (implicit), tended to be of the former type. Themes related to patterns of shared meaning across the data which were underpinned by a central idea or concept, and were relevant to addressing the research objective. Steps 1-5 were used iteratively after each set of two participant interviews, with the data pool growing with each cycle, and initial ideas open to a continual shifting throughout the process, until, after the final participant, all the data could be considered as a whole, and final themes could be confirmed. A full table of themes and related codes, evidence, and reflective comments can be found in Appendix O.

Construction of the Narrative

The fourth quality stipulation from Feldman (2007) is a clear and detailed description of how the narrative was constructed. The narrative aspect of AR is an important one, with many authors alluding to the pivotal role attributed to the 'telling of the story' (McAteer, 2013), using 'layers' of narrative (Bold, 2012), and even using the story itself, in the form of allegories, as an instrument of analysis (Dom & Willockx, 2022). Indeed, the construction of the narrative is difficult to unravel from the process of data analysis described above, and should therefore be considered alongside it. For the current project, steps 1 to 5 of the Thematic Analysis process laid out by Braun and Clarke (2006) were undertaken for each cycle of participants once they had completed the post-intervention discussion, with themes revisited when new data from each set of participants was added, before a final table of themes was created (Appendix O). These themes were then interwoven within the layers of narrative (as opposed to structuring the findings and discussion around them), so that the linear narrative of the change process could be retained.

Thus, for the first layer of narrative, a simple descriptive account using non-technical language (as endorsed by Davis, 2007, p.194) was created, detailing the feedback received from parent/co-researchers across multiple perspectives (dialectic principle) as well as how the change and decision process evolved. After this first layer of descriptive narrative was complete, a second layer of analysis and discussion was added, weaving in the reflexivity and the dialectic principles of AR (for example adding personal reflections and notes from critical friend conversations), as well as making links and seeking to explain using theory and existing literature. A final layer was then added in the form of checking against the stated quality criteria, and adding details where required - for example a missed opportunity to provide a different perspective, critique an explanation, or correct an abuse of narrative privilege. Whilst the layered approach used for the current project is supported as an effective and reliable approach to writing up (Bold, 2012), it should be acknowledge that published examples of Action Research demonstrate a wide variety of ways to present work (for example Dom & Willockx, 2024, presented four separate allegories, followed by a discussion, whereas Perez-Granados et al., 2022, presented a case study with a combined results and discussion section) – and thus there is no one accepted format or structure.

Incorporation of Multiple Perspectives

As previously noted, the quality criteria put forward by Feldman (2007) are based on those proposed by earlier researchers (e.g., Heikkinen et al., 2007; Winter, 2002), and are somewhat difficult to unravel from each other. It is unclear why Feldman (2007) included this fifth criterion as separate

from the dialectic principle considered above, but one possible explanation is that he wished to emphasise an additional requirement for researchers to 'seek other ways to represent the same data and to use them to critique the views that one owns' (Feldman, 2003, as cited in Feldman, 2007). In the current project, this was achieved via the critical friend relationship, as described above.

Explanation or Theory Explaining the Link between Actions and Outcomes

The final quality indicator or criterion put forward for consideration by Feldman (2007) is that of theory and explanation. By incorporating plausible theoretical explanations for why particular actions may have led to particular outcomes, and critically evaluating such explanations, then the validity of any claims is enhanced. The current report addressed this criterion by doing exactly that; critically evaluated explanations were woven into the discussion wherever it was appropriate to do so.

Summary of Feldman (2007) Quality Criteria

In summary, the quality criteria suggested by Feldman (2007), and which this AR project held itself accountable to, are as follows: consideration of the principles of dialectics and reflexivity, as well as how these have been operationalised; a clear and detailed description of how and why data has been collected; a similar level of description regarding how the narrative was constructed; the incorporation of multiple perspectives (used critically and reflectively as per the principles of dialectics and reflexivity); and finally, an explanation or theory explaining the link between actions and outcomes. These listed criteria offer guidance in how to approach AR, as well as a way of evaluating its quality, and a means of justifying the decisions and approaches taken. For each criterion above, an explanation of how this applied to the current AR project has been given, and will be revisited as part of the discussion in the following chapter.

Suitability of Action Research for the Current Project

AR is particularly suited to this phase of the project, with major aspects of the rationale for its use here including its flexible and practical approach, and the inclusion of parents/co-researchers and the researcher themselves as important stakeholders. These aspects are discussed in more detail below.

Flexibility

AR is a flexible approach (Dewey, 1910) offering a 'repertoire of methods and strategies' (Feldman et al., 2018, p.14). This, coupled with its focus on practice and problem-solving (Burgess & Goldman, 2006), and 'everyday human concerns' (Coghlan & Shani, 2021, p.465), aligns itself with the pragmatic research paradigm set out for the Phase One of this thesis (Greenwood, 2007, as cited in Boulton & Preiser, 2021). For example, in Walmsley and Mannan (2009), AR was seen as particularly effective at making tangible improvements for families with caring responsibilities by focusing on practical ideas and focusing on the challenges presented by the situation. Methods can be tailored to what is achievable, with decisions driven by a sensible ratio of efforts to results (Feldman et al., 2018, p.14). Furthermore, the stated aim of the current phase of this research is resource development, as opposed to testing, and is thus exploratory in nature and suited to the conjecture, tentative answers and exploration suggested by Coghlan and Shani (2021, p.467).

AR is also claimed to be particularly suited to complex social situations (Riding et al., 1995), as it allows the complexities to be considered in a flexible way and from multiple perspectives; it encourages the researcher to approach the data with an open mind, embrace uncertainty and challenge assumptions (McAteer, 2013). This is important, as the ultimate goal is a resource that works for children (and therefore parents too), and so pre-conceived notions may need to be challenged if these go against achieving such a goal. AR also lends itself to methods which are experiential, in terms of the knowledge gained through experience (in this case experience of the audiobook). Such experience and knowledge can then be reflected upon and used to decide upon future actions (Boulton & Preiser, 2021). Furthermore, AR's flexible approach does not preclude the use of other complementing methodologies - indeed the starting point of this AR project is itself built upon quantitative findings from Phase One.

Parents and Researcher as Stakeholders

Secondly, the AR approach considers participants as co-researchers with essential knowledge to contribute to the process (Boulton & Preiser, 2021). It therefore seems entirely logical that where such participants are key stakeholders (as here, and in AR in general), then they are ideally placed to provide ecologically valid insights into how an intervention might be improved, either through direct experience or as part of a hypothetical discussion. Many interventions that are both theoretically grounded and robustly designed and tested may still prove ineffective when tested in their intended 'real-world' context, and although the explanations for this are likely to be multiple and complex, parent-related reasons will almost certainly play a role. Ultimately, the opportunities a child receives, particularly in the pre-school years, will be defined by the parents and their opinions, motivations, values and beliefs, and hence parental input regarding the sort of intervention which could work well is essential.

Involving parents as co-researchers can also shift the balance of power, with the importance of shared power and equality in contribution noted earlier in this chapter (McArdle, 2004). In Loizou (2013), the use of AR was found to be empowering for parents in the appraisal of a parenting programme, in that it gave them the 'space and flexibility to be agents of their own learning and development' (p.88). Kilgour and Fleming (2001) agreed that AR facilitated a liberating and collaborative approach (again with respect to a parental education programme) that both empowered and improved self-esteem for the mothers involved. As an approach, AR respects the role of parents, and the cultural context within which they are working (Brydon-Miller et al., 2003), and policymakers, teachers and anyone else with decision-making power should appreciate this and allow for such a context to feature heavily in such decisions.

In a related point, the issue of implementation fidelity, in terms of how closely an intervention is delivered as intended by the designers, and the difficulties parents may have in adhering to such intentions (as well as the difficulties researchers have in knowing whether or not they had been adhered to) was discussed in the context of reviewing parent-led interventions in Chapter 7. Variations in parents' knowledge, skill, time available and much more mean that interventions intended to be carried out in a particular way, with fidelity, may not in fact be carried out as intended – and there is an underlying assumption that this means such interventions will not be as effective. However, fidelity is not consistently related to outcomes (Trickett et al., 2020), and parents may have different ways of implementing an intervention which are more effective than the intervention designer originally

conceived – providing further support for parental feedback and the use of AR methodology for the current project.

Having discussed the importance of parental input for the development of an intervention resource intended to improve children's vocabulary, it is equally important to discuss the researcher's own position as a stakeholder and collaborator in this AR. As noted earlier in this chapter, the term 'second-person practice action research' has been applied to this project, in that the topical issue is one of 'mutual concern' (Coghlan & Shani, 2021, p.468). Children's vocabulary, and more broadly their language development and overall well-being, plus related issues of fairness and equality, could well be considered an issue of 'mutual concern' for many people. For the researcher, as a parent, and as a person with a background of working with primary school-aged children and a current occupation of completing a PhD and teaching within the area of children's language and overall development, the outcomes of this AR project are particularly important on both a personal and professional level. For these reasons, the researcher positions themselves not only as a collaborator, but also a stakeholder, in a similar approach to that taken by Kwiatkowski (2019). As a stakeholder who is responsible for the initiation of this project, and it's eventual write up, due acknowledgement must be given to how the researcher's own values and experiences, particularly those relating to teaching children and of being a parent, and what has 'worked well', will necessarily influence the whole AR process, including the collection of data, its interpretation, and the construction of the narrative/write-up.

A Model of Action Research

Numerous models of AR exist. In the pursuit of simplicity, this project will loosely follow that of Riding et al. (1995), which involves utilising the plan – act – observe – reflect cycle, with each cycle feeding into the next (see Figure 1). Despite the intended simplicity however, due acknowledgement should be given to the unpredictability, non-linearity and 'messiness' of Action Research projects, as described by McNiff and Whitehead (2002). AR projects should seek to embrace these facets of human behaviour and experience, and continually review and readjust the process, whilst aiming to communicate the 'seeming incoherence of the process in a coherent way,' (McNiff & Whitehead, 2002, p.12).

Limitations of using Action Research

AR as a methodology is increasingly associated with professional development within an educational context, and about improving one's own practice, which differs from the way it is being employed for this project. Nonetheless, AR can be used across a wide range of contexts and for different purposes (Bradbury, 2015; Banks & Brydon-Miller, 2019). Indeed, Kurt Lewin, the social psychologist who is thought to have first coined the term Action Research (McKernan, 1991), employed the methodology in his own work concerning people affected by post-war social problems. Furthermore, the current AR has been defined as 'second-person practice action research,' a form of AR which refers to collaborative co-inquiry and shared action with others on issues of mutual concern (Coghlan & Shani, 2021, p.468), with both the parent participants/co-researchers and the researcher positioned as stakeholders.

A more general limitation associated with AR is lack of rigour, largely due to the (usually) qualitative nature of AR, and the traditional concerns of validity, reliability, replicability, objectivity

and generalisability become less clear. Indeed, AR actively encourages subjective researcher involvement, and projects are often small scale and limited in scope, so are not seeking to be generalisable (or claiming to be). The framework of quality criteria from Feldman (2007) which have been clearly outlined above, including an explanation of how each one applies in the current project, go some way in circumventing this limitation. This explanation included a consideration of the ethical aspects of research, for example the potential for abuse of narrative privilege. In a related point, the current project received a favourable ethics opinion from the NTU School of Social Sciences Ethics Committee before commencement.

A further potential limitation of using an AR approach, connected to the points made above regarding quality criteria and particularly narrative privilege, are the difficulties associated with writing up the project. At this stage in the process, the flexibility inherent in the AR approach could be considered both benefit and hindrance, in that there is no generally agreed framework for writing up AR, and thus the challenges of producing a coherent linear narrative detailing the change process, whilst simultaneously providing additional layers of analysis and discussion related to the themes identified via the chosen analytical process (not necessary linear in nature) – is not straightforward. Indeed, depending on how the write up is structured, the cyclical nature of AR poses a challenge to linear story-telling due to the 'ongoing and conversational relationship between data and literature' (McAteer, 2013, p.140). Nonetheless, proponents of AR suggest embracing the messiness, unpredictability and non-linearity of AR projects (McNiff & Whitehead, 2002), and to try and communicate this in a coherent but flexible way – indeed such flexibility allows for multiple approaches regarding structure, as evidenced by the variety of published Action Research reports available. McAteer (2013) suggests a further challenge in achieving a balance between authenticity and appropriate academic tone in a written report, exacerbated by the messy and cyclical nature of AR, and the central position of personal voice, and suggested that intended audience should play a role when assessing how this balance should be weighted. To address some of the other challenges associated with writing up AR, the process of constructing the narrative has been considered above in detail within the framework of the quality criteria. In summary, for the current project, the narrative aspect was emphasised via the initial descriptive layer of capturing the feedback and change process in simple non-technical language (Davis, 2007), followed by layer of analysis and discussion, and then a final layer involving further consideration of the quality criteria.

Summary of Chapter 8

This chapter has outlined the methodological approach for Phase 2 of this thesis, which aims to explore how a vocabulary intervention can be effectively incorporated into an audio format using an Action Research approach seeking the collaboration and feedback of parent participants as co-stakeholders alongside the researcher. AR is a flexible approach involving iterative cycles of planning, acting, observing and reflecting (as depicted in Figure 1), and focused on creating practical, actionable knowledge. From an epistemological stance, an argument was presented for a 'middle ground constructivist' position on what can be considered knowledge (Feldman, 2007), with some claims to knowledge accepted at face value, and others requiring evidence and consensus around whether the claim in question does indeed constitute knowledge and truth. In conjunction with this position, the framework of quality criteria for AR proposed by Feldman (2007) were each considered in turn, which included the principles of dialectics and reflexivity, as well as how these have been operationalised; a

clear and detailed description of how and why data has been collected; a similar level of description regarding how the narrative was constructed; the incorporation of multiple perspectives (used critically and reflectively as per the principles of dialectics and reflexivity); and finally, an explanation or theory explaining the link between actions and outcomes. Within this framework, a range of key methodological issues were considered, including details of sample and resources, and methods of data collection, analysis and construction of the written report. Finally, the suitability of AR for the current project was considered, alongside potential limitations.

Chapter 9: Findings and Discussion (Phase Two)

The aim of this Action Research was to explore how effectively a vocabulary intervention (via exposing children to more verbs) could be incorporated into a simple audiobook format to be delivered by parents, in this case the fairy tale *Goldilocks and the Three Bears*. In Phase One of this thesis, verbs were shown to be consistently predictive of a child's overall vocabulary diversity, suggesting that the learning of verbs specifically may help to drive children's language acquisition in general, and be particularly important for children with smaller vocabularies. To this end, and as explained in detail in Chapter 8, parent-child participant dyads were recruited two at a time and provided with one of two versions of a short audiobook resource. To reflect how a future efficacy trial would look, two versions of the audiobook were created, one with a narrower selection of repeated verbs, and the other with a more challenging selection of non-repeated verbs. Participant dyads were assigned in order of recruitment to either version, and parents were asked to play the audio resource to their child at least once per day over a one-week period.

As noted in the previous chapter, the following account has been constructed via a three-layered process, which is intended to combine a mostly linear narrative description of the findings, an analysis and discussion incorporating literature and theoretical explanations, and an addressing of the quality criteria described in the previous chapter. Together, these layers will attempt to capture the change process. At the head of the account, the themes which emerged from the Thematic Analysis are presented in a table alongside a theme description and example evidence (for a full table of themes and related codes, evidence, and reflective comments see Appendix O). These themes will then be interwoven within the layers of narrative (as opposed to structuring the findings and discussion around them), so that the linear narrative of the change process can be retained. The benefits and limitations of structuring the write up in this way will be considered at the end of this chapter.

Table 10

Summary of Themes with Theme Descriptors and Example Evidence

Theme	Theme Description	Example Evidence
1. Existing Routines	Anything which related to a family's existing routine, including (but not limited to) shared reading, experience of audio and parents' expectations.	*Normal routine includes lots of traditional reading, and also audiobooks: at bedtime she [Emily] will often choose an extra audio after reading – but not every day, and there is a wider choice.* *Nathan will 'read' on his own.*
2. Ease of Audio	Anything which related to the facilitation of the audiobook in terms of, for example it's portability and convenience.	*Easy to administer. Played in different settings e.g., bath, car.*

		Hope would often decide when it would be played.
3. Engagement and Learning	Anything which related to perceptions of how children engaged with and learned from the audio.	*First time was very interested – perhaps due to the novelty of being allowed to touch phone (isn't normally).* *In the first few days, Hope would remember bits of the story and talk about them without being prompted.*

Cycle 1: the Unadapted Audiobook

The initial version of the audiobook given to the first round of participants was an unadapted audio recording of *Goldilocks and the Three Bears*. Of the two versions, Sam (mum) and Hope (child) received the simpler version, and Eloise (mum) and Joseph (child) received the more complex version.

Sam and Hope

Hope was allocated the simple version of the audiobook, and Sam thought that the length and ability level was about right, but wondered whether families with lower literacy levels or those who did not speak English as their first language might experience more difficulties. After Sam was made aware that this was the 'easier' version of the audiobook, as randomly allocated, she provided some interesting feedback about the importance of repetition within the story, in that she felt that the repetitive element of the story (which was more enhanced in the easier version of the audiobook) had a positive influence on Hope's engagement, as well as making her more likely to retain knowledge of the words that had been repeated. The more complex version of the audiobook would necessarily lose some of this repetition due to being more verbally diverse, and this could have unintended consequences.

> *Notes: This is important, as the harder more verbally diverse one will not be quite as repetitive, and this may detract from its positive influences on language…?*

Research indicates that word frequency is important for word learning, in that the more frequently a word appears in the language (the more times a child hears it) the more likely they are to acquire it – although the effects of word frequency may be qualified by other factors, as discussed in Chapter 3. Increasing word frequency by repeating words within a story, or repeating the same story over and again has also been shown to be effective for word learning. In Horst et al., (2011), 3-year-old children learned more novel words from shared storybook reading when the words were contained in three storybooks that were repeatedly read within the same session, as opposed to when the (same) words were contained within nine different storybooks. Here, the words in question were novel nonsense nouns, and the children in the repeated storybook reading condition were able to both recall and retain novel words more accurately when compared to children in the different storybook condition, who could only recall the novel words on some occasions and had little evidence

of retention. Whilst this study supports the idea of repetition being helpful for word-learning, and specifically the idea of repetition within the same familiar story, there are limitations in generalising the results to audiobooks, in that part of the explanation provided for children's more effective learning of novel nouns from repeated exposure from the same storybook was that the repeated exposure was to both the word and an associated referent (in the form of an illustration) simultaneously – a characteristic which is necessarily missing from audiobooks.

Whether or not listening to familiar audiobooks containing repetitions of new words over and again would be more effective for children's learning than listening to a variety of different audiobooks with the same words clearly requires further research, and yet regardless of this outcome, repetition, (particularly when combined with rhyme) adds a rhythmic, nursery rhyme type quality to a story. Such a quality appears to be both highly engaging for children (and the adults who read with them) – as evidenced by the numerous children's authors, including the hugely popular Julia Donaldson, who have adopted this strategy, but also beneficial for their language outcomes, in that exposure to nursery rhymes in early childhood has been repeatedly associated with positive language outcomes for children (Harper, 2011). Therefore, in light of the likely benefits of repetition for children's word learning, it is important to bear in mind the implications for the aims of this intervention, and the theoretical starting point that verbs are driving vocabulary acquisition (see discussion in Chapter 6). Clearly care is needed in order to strike a balance between achieving the diversity of language desired without losing that sense of repetition, rhythm and rhyme which appears to engage young children, and potentially help them learn new words.

Hope listened to the audiobook every day, or more sometimes, often asking for it to be played without prompting. Sam expressed surprise at how well Hope engaged with the story, as despite doing a lot of traditional reading at home, she had never listened to an audio before (nor did she appear to know the story of *Goldilocks and the Three Bears*). Sam reported that the audiobook was easy to administer (as was intended, see rationale for using audiobooks in Chapter 7), of a high quality, and lent itself to being played in different contexts:

> *Notes: Played in the car, but mostly at home, in different rooms.*
>
> *Notes: She didn't appear to know the story already, and asked mum why they didn't have the book.*
>
> *Notes: S was surprised with how well she engaged – 90% of the time engaged – and didn't have to be encouraged to listen.*

Importantly, Hope would remember bits of the story outside of listening time and talk about them without being prompted. She also employed her own toys as make-shift props that were linked to the audiobook, and Sam felt this additional physical dimension complemented the audio resource, and provided another positive influence on engagement.

> *Notes: Hope: "I like that she eats the bear's food it's a bit funny."*
>
> *Notes: H found a dolly and S asked "Is that Goldilocks?" after which the doll became Goldilocks all the time.*

Sam's feedback about the audiobook was extremely positive, and interestingly she didn't think any adaptations to the content or format were necessary, although she agreed that Hope would

be delighted to listen to a story where she was the named main character. Conversely, Sam hypothetically noted an uncertainty about how well Hope's older brother would have engaged with the audio mode, and attributed this to personality differences between the two.

Eloise and Joseph

Eloise and Joseph were allocated the more complex version of the audiobook, and Eloise felt that the story was pitched at about the right level, with a brief chat required about the meaning of the word 'angelic.' The fact that Eloise took the time to explain an unfamiliar word, perhaps in a similar way as she would have done when sharing a physical book, was interesting to consider both logistically (was this done afterwards, or as the audio was playing?) and in terms of what it implied about the family's responsive approach to learning. Eloise agreed with Sam in that the quality was good and the resource was easy to administer, demonstrated by the different settings in which it was played. Eloise thought that Joseph's level of engagement was good, with the added caveat that staying and listening wasn't always the best marker of engagement, as often there wasn't a choice (e.g., in the car). As the week progressed, Joseph began to talk about the story (one he was already familiar with) outside of listening times, and sometimes to copy the voices.

> *Notes: Joseph: "I heard one where they put porridge on Goldilocks' head!"*
>
> *Notes: Joseph: "I like the big bear!"*

In considering how the resource might be improved, Eloise wasn't initially convinced how effective personalising elements of the story would be, with respect to familiar names or narrator, but she did accept that this could be important for some children who felt more marginalised. Eloise suggested that adaptations to the content or overall genre might be important so that choices could be tailored to individual preferences. Joseph evidently took to repeating the different voices, and thus it is perhaps worth considering that this part of the audio is particularly salient to children, and to include more target vocabulary within speech, and more speech overall, might be an effective strategy to ensure engagement.

> *Notes: J was particularly interested in the gory 'scary' stuff (he would talk about this at other times) e.g., gobble you up, and the voices.*
>
> *Notes: …mostly he is quiet. But as soon as bear voices start, he becomes animated and joins in and appears to be more actively engaged.*

Eloise pointed out that Joseph enjoyed an established reading routine with his parents, and felt that his engagement with traditional shared reading far surpassed his engagement with the audiobook here, as measured by his level of interaction: asking lots of questions, pointing at things, etc. She felt that something physical to hold or look at – such as a storyboard or related props, had the potential to elevate that level of interaction. A related physical prop might also serve to channel any extra energy back towards the story by encouraging the re-enactment of parts of the story. This links in with the discussion in Chapter 7 around the importance of contingent talk for vocabulary learning, and how the presence of accompanying physical objects, such as a toy character or a paired book with pictures could serve as a focus of contingent interactions. The talk of 'animation' and 'joining in' is relevant to the idea of experiential learning, also shown to be important for children's learning (see Duncan et al., 2019): as noted in Chapter 7, the audiobook format does not have inherent

experiential qualities, however actions which take place alongside the audiobook, facilitated by physical props, parents or otherwise, have the potential to create a learning experience which is less passive and more experiential.

Action Taken (1): Introduction of Accompanying Physical Resources (Finger Puppets)

Sam and Eloise provided a wealth of feedback across the three themes of existing routines, ease of audio (facilitation) and engagement and learning. Whilst there appeared to be some agreement with respect to the convenience and ease of audiobooks, and a generally good level of engagement reported, there were clearly areas in which the audiobook was lacking, or had the potential to be improved. Whilst beyond the scope of this small-scale AR project, providing a choice of audiobooks was noted as an important characteristic if individual preferences were to be accommodated, although it was also noted that engagement with the format in general could differ between children (using siblings as a hypothetical example). An important issue was raised with respect to the design of the two versions of the audiobook, in that care needed to be taken to ensure balance between providing a diverse range of vocabulary (verbs in this case), but also repetition, which alongside rhyme, is also beneficial for children's language learning and overall engagement with reading and listening. Existing routines, particularly around reading, were expressly stated or implied, and confirmed strong existing shared reading routines. In connection with these existing routines, both Sam and Eloise noted the presence of extra-textual talk and the importance of physical, concrete things – and ultimately the potential of the latter to elicit the former (as well as improve overall engagement).

The presence of unprompted, extra-textual talk outside of listening times is an effective defence against one of the criticisms of audiobooks: their inability to facilitate interaction between parent and child. Due to the tendency to play audiobooks continuously, and their capability of being played without an adult present, it can be argued that they lack the natural opportunities for questions, answers and decontextualised talk that traditional shared reading provides, particularly when done in a dialogic format (Pillinger & Wood, 2014). Whilst the extra-textual talk and resulting interactions detailed by the participants here were not always during the listening experience, so arguably less 'contingent' in nature, the audio still appears capable of eliciting extra-textual talk between parent and child - and this is important seeing as such interactions are deemed to be the crux of where children learn language (Goldin-Meadow et al., 2014; Tamis-Lemonda et al., 2014), as well as being associated with greater enjoyment in reading later in childhood (Baker et al., 2001). Talk which happens outside of listening times is arguably more decontextualised in nature, in that it is talking about something which is removed from the here and now – yet, as noted in Chapter 7, this type of talk is also important for children's language learning, particularly during the 4th year of life (see Rowe, 2012, p.1771).

How could the audio format harness more expressive talk from children, and therefore the parent-child interactions that often follow? Children often talk about what they are attending to in that very moment – referred to as contingent talk (McGillion et al., 2017) – and for young children, that something is often physical in nature, such as a toy or household object. The Emergentist Coalition Model of word learning (Hollich et al., 2000) discussed in Chapter 1, supports the idea that children use a variety of cues to help them learn words, with one of these cues being perceptual salience, (although the use of different cues may depend on both age and context). In other words, if a child is

playing with a physical object, then such an object is likely to be highly salient to them, and contingent talk related to such an object is an optimal context in which to acquire new words (Akhtar et al., 1991).

The potential of linked physical resources to facilitate extra-textual talk and further interaction around the audiobook can be related to another line of discussion: the central role of verbs. On consideration of the aforementioned theoretical starting point that verbs are driving vocabulary acquisition, Eloise pondered over whether orally exaggerating the verbs within the audiobook might make them more salient to the child and therefore improve the probability of them being remembered. This intriguing suggestion was duly explored, but in practice the effect proved difficult to achieve whilst still retaining the rhythmic quality of normal speech. The suggestion did give rise to an alternative line of thinking however: emphasising the verbs did not necessarily need to be done orally. For example, when Hope named the doll *Goldilocks* and employed her in acting along with story, those very actions were the verbs in the story, and by re-enacting them, she was simply emphasising the verbs in an alternative way, as well as creating a more experiential and active learning experience. This is a similar idea to the moving and story-telling intervention tested by Duncan et al. (2019), where a 12-week intervention which involved a combination of language and movement was found to significantly influence children's expressive language, with both aspects of the intervention connected by a popular children's book.

Beyond the idea that linked physical resources might aid children's word learning via engagement, contingency and facilitating verb learning, as described above, the fact that both Sam and Eloise themselves felt strongly about the importance of concrete and physical aspect of learning resources added additional weight to the argument for including some sort of physical accompaniment alongside the audiobook. The change was also pragmatic in nature in that it could be easily achieved from a cost and implementation perspective. The first change to be applied after the initial round of participants was thus to introduce accompanying physical props in the form of four finger puppet dolls: Goldilocks plus three bears. The intended outcome of this action was to help children to visualise and engage with the story, and boost word learning by facilitating more extra-textual talk about the story, potentially leading to more conversational interactions, as well as making the situation more contingent in nature and providing the means for a subtle emphasis on verbs (by using the dolls to re-enact the various scenes). The adaptation also aligned with parents' opinions regarding the importance of the physical, concrete aspect of learning environments. Whilst the inclusion of additional resources adds some complexity to the simple, easy facilitation of the audiobook, the dolls will be provided with the caveat that they are optional and can be used at any time, in whichever way the child wishes.

Cycle 2: the Finger Puppets

The second version of the audiobook given to the second round of participants included the same audio recording, but this time was accompanied by the finger puppets. Pippa (mum) and Tom (child) received the simpler version with less verb diversity, and Cara (mum) and Leon (child) received the more complex version with a greater variety of verbs.

Pippa and Tom

Pippa thought that the length, difficulty level and quality of the audiobook was appropriate. Tom was familiar with the audiobook format, and would often listen in the car, sometimes with an accompanying book. Tom was played the audiobook four times over the course of the week, and Pippa admitted that ring-fencing dedicated listening time (despite the audio only lasting approximately 3 minutes) was difficult set against the context of being a busy mum of four children with full time work and additional commitments. A further technical barrier presented itself by way of the audio needing to be played on a static and immoveable device attached to one room (rather than a smartphone), and thus negated one of the main intentions concerning ease of facilitation.

Notes: Would sometimes only get home as children were being put to bed.

Two of the most important advantages of the audio format are accessibility and convenience, and this was corroborated by participants in the first cycle who discussed the ease of playing the audiobook in different rooms in the house, in the car, and whilst other activities were happening simultaneously. Children could also be allowed some level of independence in deciding when and how often the audiobook could be played, although this depended on whether they were permitted access to their parent's phone, or indeed whether the phone was available and not being utilised for other tasks. These intended advantages aligned with the idea that parents could be nudged (Thaler & Sunstein, 2008) into desirable habits (desirable as they would benefit their children's language development) when an effortless option was provided.

Pippa, on the other hand, offered an alternative perspective, demonstrating how a number of simple obstacles, particularly when in combination, can seriously impede the ability for parents to facilitate interventions in the home. This is particularly true if the intervention is difficult and time-consuming, however it is also true (as here) when the intervention is designed around ease of delivery. In this case the obstacles were time pressures and technical problems, but they could be anything, and are likely to be compounded where multiple challenges are present. Consideration is thus needed to ensure such barriers are mitigated against as far as possible, for example by providing the audiobook via different platforms as opposed to via email alone.

Another factor that may influence ease of facilitation and children's language development overall is the presence of other siblings. Such a presence is often a helpful thing, and Pippa discussed the positive role Tom's siblings played by listening to the audiobook alongside him, encouraging him to stay listening, and treating the whole event as a communal family affair.

Notes: Brother and sisters also listening so would interact with them.

On the other hand, one can foresee a situation whereby siblings, particularly those of a different age, have the potential to distract the listening child. Families with larger numbers of children in their household are associated with poorer language outcomes, in terms of vocabulary or other language-related skills (Laing & Bergelson, 2024; Muñez et al., 2022), particularly for younger siblings (Hoff-Ginsberg, 1998), possibly explained by the resource dilution model (Blake, 1981). This issue is obviously not confined to audiobooks – but any kind of intervention or interaction that could benefit language outcomes, including traditional shared reading – however it is nonetheless worth bearing in mind when considering the context and timing that is optimal for listening.

Tom liked playing with the finger puppets of Goldilocks and the three individual bears, and they would be present in Tom's play both during listening and at other times. Pippa felt that the props helped Tom to visualise the story, although she noted that often his re-enactions with the puppets did not match the story, and belonged to imaginary play or a different story. This suggests that, for Tom anyway, the finger puppets may not necessarily have had the intended effect of focusing attention, and any in-the-moment contingent or subsequent decontextualised interactions, on the audiobook content and the verbs contained within it – although they may of course facilitated other beneficial interactions and imaginative play. As this Action Research study is exploratory in nature, it is difficult to establish the exact extent to which Tom (or any of the children) was engaging with the audiobook content in a passive or active way – and yet the observations are a useful starting point in the consideration of how such a resource may work in practice, and how it could be developed.

> *Notes: Used props but not necessarily along with the story… wouldn't always listen to whole story, but would role play a different story.*

The conversations around related physical entities moved onto audiobooks that came with an accompanying physical book that could be read whilst listening. Pippa was unsure about this, as her experience of using such resources with her own children in the car had suggested that they did not use the two resources simultaneously. In general, she felt that it was important for children to develop their own imaginations and visualise stories heard over audio themselves – with the caveat that young children may benefit from physical props. These thoughts are somewhat contrary to the evidence presented in Olson et al. (2024), where audiobooks with accompanying text were found to significantly and positively influence vocabulary, albeit for older children (mean age = 9.47 years). On the other hand, pre-schooler books usually have more pictures than text, and so it is the pictures and how they influence the child's visualisation of the story (rather than the words themselves) that are likely to be the focus of an accompanying book. As Pippa commented, visualisation of the story did not necessarily need to be prescribed via pictures in a book – but could be achieved by children utilising their imaginations.

Pippa was also undecided about the idea of introducing familiarity into the story, either by content or delivery. She concluded that familiar names and places might be worth trying, as it might focus the child's attention and result in them listening to a higher proportion of the words than they might have done otherwise.

> *Notes: Own name and familiar place names might be worth a try, even if it just stole their attention for a few extra moments, meaning that they listened to a higher percentage of the story than they would have done otherwise.*

Cara and Leon

Leon was allocated the harder version of the audiobook, and Cara thought that the length and difficulty level were about right, and that the audiobook was easy to facilitate by just pressing play on her phone: they played the story once every day, on average. Cara reported Leon's level of engagement with the audiobook to be good, which might be partly explained by his familiarity with the format (he owned his own Yoto player with a variety of cassettes that he could independently insert and play). Leon also had an established shared reading routine in place, but was not familiar with the story of *Goldilocks and the Three Bears.*

Notes: More engaged than mum thought he would be. Would stay in the room and listen (but some of this was possibly due to mum's and sister's encouragement).

Notes: He also normally engages well with traditional storybooks.

Existing routines, particularly those around shared reading, are important to mention here because they are somewhat difficult to disentangle from the story of the child's engagement with the audiobook. Where children are already familiar with physical books and shared reading with an adult, this suggests that they will be similarly capable of listening to and engaging with stories in a different format (audio) due to well-developed listening and attention skills, and knowing what is expected of them. Thus, generally good levels of engagement with the audiobook might simply be a product of existing routines, and because of a certain level of obedience within the parent-child relationship: the audiobook was presented by the parent as something desirable to do, and listening was encouraged. If children are missing prior experience of reading, or an authoritative parent-child relationship (or both), then engagement with the audiobook might prove more of a challenge.

In an opinion similar to that of Pippa's, Cara thought that the finger puppet props probably helped Leon to visualise the story more effectively (although this was difficult to ascertain one way or another), and certainly elicited extra-textual talk about the story when he was playing with them outside of listening times.

Notes: Use of props… was effective in that mum thought they helped L to visualise the story – he carried them round with him all week, which also seemed to facilitate talk about the story outside the story listening times.

The physical presence of the puppets also prompted both parent and child to think about and remember the story – particularly when the toys were carried around with the child consistently throughout the week, as new toys often are. In this way, the toys acted as a focal point for both parent and child, facilitating contingent interactions on the topic of the audiobook, either in the moment, or in a more decontextualised way at other times. Whilst the dolls were not necessarily utilised for their intended purpose (e.g., they were not always used alongside the story) it could be argued that the outcome was still beneficial (and certainly not detrimental) as the dolls prompted the child to play and talk in an action-focused way, with both contingent talk (Akhtar et al., 1991) and verb usage (Horvath et al., 2022) associated with positive language outcomes for children.

When the discussion moved onto the idea of introducing elements of familiarity to make the story more engaging and easier to identify with, Cara considered that this might be effective, and drew a comparison with a book that Leon had enjoyed which had incorporated his own name as one of the central characters. In terms of a familiar voice or accent, Cara admitted that she had not noticed the narrator's accent, only that the quality of the audio was good, and that the narrator's voice was as she had expected. The conversation then evolved into thinking about children and parents with different backgrounds (perhaps those with stronger regional accents or from a minoritised ethnic background), and Cara pondered over whether they would have felt differently about the voice and accent of the narrator, but wasn't sure. She suggested the possibility of employing a child to narrate the story.

Notes: …thought quality of audio was good, but didn't notice the accent, in that she thought it sounded professional and normal – as an audiobook should sound.

Action Taken (2): Introducing Elements of Familiarity

The feedback provided by Pippa and Cara added to the existing data pool provided by Eloise and Sam. Many of their comments corresponded with those provided by the earlier participants/co-researchers, for example around the theme of existing routines, where both Pippa and Cara discussed their existing routines around shared reading and audio, with Cara attributing Leon's engagement to this familiarity with the audio format. Both Pippa and Cara discussed the influence of the finger puppets in a positive way, in terms of their child's desire to play with them, and on some occasions, their role in eliciting extra-textual talk about the story – however both agreed that it was not always easy to know if the finger puppets were helping their child to listen and engage with the story. Related to the ease of audio theme, a contrasting perspective was noted by Pippa with respect to the potential barriers to facilitating the audio, and how even an intervention designed to be convenient can still be difficult to administer for those who are time-poor. Finally, both Pippa and Cara considered the concept of identifying with a story, through its characters or otherwise, an important one, but it was not necessarily something that they had thought about in relation to their own children. Both parents considered themselves to be in a relatively fortunate position in terms of financial security and access to resources, and did not initially notice aspects of the resource that others might have, for example the narrator's voice or how different children might identify with the specific elements of the content. Despite this, they were keen to discuss these issues in relation to other hypothetical groups once prompted.

Variation of accent was one possibility that was discussed as a potentially important feature for engaging under-represented groups. Both Pippa and Cara had not noticed the voice used to narrate the audiobook (male, neutral accent), quite possibly because this was the style of voice to which they were accustomed and expected in this context. Had the voice been something out of the ordinary – less mainstream or neutrally accented – their attention might have been aroused towards this particular feature. Cara contemplated the potential importance of accent in a hypothetical sense for children and parents from different backgrounds (perhaps those with stronger regional accents or from a minoritised ethnic background), but was undecided on the impact this would have. Some vocal characteristics, such as the use of motherese with young children, whereby speech is simpler, higher pitched and with exaggerated intonation, have been clearly and positively linked to children's language acquisition (Fernald, 1985), but there is only a limited amount of research on the effects of other vocal characteristics, such as how different accents or ages of voice may influence children's engagement and learning (Cara also suggested using a child's voice to narrate the story), and whether such effects may be more important for some groups of children and not others. On the other hand, research has indicated that up to 50% of children struggle to identify with the characters in books they read, making them less likely to engage with and enjoy reading. This was more likely to be the case for children on FSM, or from a non-white ethnic background (Cole et al., 2022). Providing audiobooks narrated by a range of accents might be one way to help children identify with a story, and this may be particularly true for groups who tend to be under-represented in the general media, with voices and accents that are therefore less heard (Lee, 2019).

Another way to help children identify with stories they read or listen to is to change the names or characteristics of the features in the story, for example the main character's name or their appearance, or the setting where the story takes place. Again, as with accent, this could be particularly

relevant to under-represented groups who often feel unable to identify with the characters in unfamiliar stories (Cole et al., 2022). Audiobooks, when compared to physical picture books, perhaps go some way in mitigating against these feelings, in that the listener can infer the character's appearance, an inference that is likely to be closer to their own image than the picture in a book. One simple way of altering the character so that they become familiar and easy to identify with is to use the child's own name, or by introducing a setting with which the child is already familiar. Whilst Pippa was not sure if using a familiar name would be beneficial, she did concede that it may focus the child's attention and result in them listening to a higher proportion of the words than they might have done otherwise. Cara supported the idea of familiar names more strongly, and again linked this support to existing routines and familiarity with the concept (Leon already had a book which incorporated his own name, and enjoyed reading it). Research has shown that children attend more readily to familiar words, particularly their own name (Bortfield et al., 2005), as these words serve as an anchor point to surrounding (potentially unfamiliar) words, easing the segmentation process – thus not only will the use of familiar names or places capture the attention and ongoing interest of the listener by making the story more relevant to them, they are also likely to assist in the segmentation process for unfamiliar words and result in more words of the story being understood.

Considering the role that interest plays here is also relevant, in that if a child's interest can be piqued by a familiar name (their own), this could be beneficial for their engagement with and learning from the audiobook. Interest can be described as content-specific engagement over time (Hidi & Renninger, 2006), although the two terms 'interest' and 'engagement' are often used inter-interchangeably, particularly in a school or home context. 'Triggered situational interest' is the first phase of Hidi and Renninger's four phase model of interest development (Hidi & Renninger, 2006) explaining how interest develops and deepens over time. During this first phase, attention may only be fleeting, and gained through the presentation of a stimulus that is incongruous, exciting or personally relevant (in this case, the child's name). This fleeting level of attention (that the child may not even be aware of) is likely to require support and encouragement to nurture into the next phase of deepening interest, but nonetheless represents an important initial engagement, and is a precursor for developing a deeper interest in particular topics or activities. For children who are not already in this initial phase of interest, a familiar name may provide this initial trigger to enter it. A similar argument may apply to parents with respect to the important role they play in their child's language development in the pre-school years. Here, the use of familiarity (their own child's name) may trigger attention over a threshold that is the difference between facilitating the audiobook and not doing so. This is another good example of how subtle adaptations to the environment (in this case, adapting particular features of the audiobook) may 'nudge' parents into making a desirable choice (Thaler & Sunstein, 2008).

Thus, after consideration of parental feedback alongside the researcher's own reflections and relevant research and theory, particularly findings indicating that many children struggle to identify with the stories on offer, and elements of familiarity may help trigger their and their parents' interest - the second change to be implemented was to incorporate simple elements of familiarity into the audiobook. These elements of familiarity would take the form of the child's name (replacing *Goldilocks*) and hometown (replacing *the woods*). The finger puppets would be kept where practicalities allowed, minus Goldilocks (due to Goldilocks' character now being named as the child). Despite the discussion with Pippa and Cara on the idea of identifying through familiarity taking on a

more hypothetical character (i.e., none felt it was particularly relevant for their own child), it was still the generally held view that this was a worthy avenue of exploration for the benefit of those children who may be less likely to identify with the story (this might be particularly relevant for minoritised groups as suggested by the National Literacy Trust's National Literacy Survey, Cole et al., 2022). The intended outcome of this change was therefore to see whether future participants felt that the elements of familiarity influenced their child's engagement with the audiobook.

Cycle 3: Familiarity

The third cycle of Action Research involved incorporating the child's name and hometown into the audiobook script, whilst retaining the finger puppets where logistics allowed. In this round, Caroline (mum) and Anita (child) were given the easier, more repetitive version of the story, and Amelia (mum) and Emily (child) the harder version.

Caroline and Anita

Caroline commented that the resource was easy to administer, and Anita was excited to listen to begin with, recognising the name of her village and using the props to act along with the story and re-tell it afterwards. Caroline considered the use of different voices as a further theme of familiarity and identification with the story, and thought it might be appropriate to investigate narrators of different ages and with different accents.

> *Notes: C thought that the familiarity of the story was a good thing… different voices might be appropriate, perhaps with some different accents and different ages as the narrating voice (but not too much of an accent as difficult to understand).*

Despite her initial excitement, Anita's engagement proceeded to fluctuate throughout the week, something which Caroline attributed to an over-familiarity with the story of *Goldilocks and the Three Bears,* and because she felt the story was both too short and too easy (this is in contrast to other parents, who generally felt the audio was pitched at around the right level, regardless of the version they received). Caroline reported that Anita enjoyed reading with an adult, had plentiful access to books in the house and had shown a keen interest in learning to read.

> *Notes: There are already lots of books in the house, and A reads a lot and wants to learn to read. She likes the interaction of sitting together with someone.*

With the exception of Caroline, the majority of parents deemed both the difficulty level and length of the audiobook to be about right. These judgments of suitability could partly be due to the influence of the fictional genre chosen (fairy tales are stories that are generally associated with being appropriate for children), but also something to do with this particular sample's broad experience of reading age-appropriate books with their children.

Pitching the audiobook at the right level in terms of difficulty and length was an important consideration, as children are understood to build their vocabularies incrementally, using the context of familiar words to infer meaning of less familiar words (e.g., via associative learning and syntactic bootstrapping, see Chapter 1). Too many unfamiliar words may impede this process, and a story that is too long is unlikely to be fully attended to by a child under the age of five who has a less developed attention span than an older child or adult. In general, children (and adults) tend to be motivated

towards learning activities that provide challenge without diminishing feelings of competence or self-efficacy (e.g., Self-Determination Theory, Deci & Ryan, 2000; Social Cognitive Theory, Bandura, 1986). Whilst one of the notable strengths of audiobooks is that they can facilitate access to a wider range of books beyond the abilities of the reader (Miller, 2010), this relates to difficulties in reading the written word, and does not suggest that the audiobooks children are exposed to should be pitched well beyond their verbal language skills. This is particularly relevant for parents who might not speak or read the English language proficiently, or who might be struggling readers themselves.

In a similar vein of thought to Sam in Cycle 1, Caroline suggested that the audiobook might have been a more effective approach with Anita's older brother, due to him having a different personality and temperament. The idea that individual differences (between children, parents and families) may influence how well a child engages with the audiobook format seems logical. Research and theory support the idea that we are usually more motivated to pursue and re-engage with activities that interest us (Hidi & Renninger, 2006), and beyond personality and temperament differences already mentioned, there is likely to be a wide range of preferences regarding story genre and content – as supported by Eloise's comment in Cycle 1 about tweaking the content to be more gruesome. A child who is interested has a positive effect via two pathways: a parent is more likely to pursue a language-related behaviour with a child that they perceived to be 'interested,' and the interested child is more likely to engage and attend to the content.

Individual differences pose an interesting problem however, in that a single audiobook - through which this Action Research sought to explore the feasibility of audiobooks in general - will inevitably fail to satisfy every individual preference, and hence comments directed at the content of the audiobook are not necessarily targeted at the audio format in general. As would be the case with physical books, providing some level of choice in practice would be imperative. It is perfectly feasible that the target verbs (or certainly a similar selection) could be incorporated into a variety of genres with varying content, and in practice, commercial providers of audiobooks have a large variety of stories and non-fiction titles available (e.g., Yoto has over 1000 titles available). Providing choice is similarly associated with higher levels of motivation to pursue an activity in that it provides a greater level of autonomy (Ryan & Deci, 2000), as well as a greater probability that something of interest will be found. Individual differences in terms of family needs and routines are also relevant; one particular genre or format of delivery may suit one child or family but not another, and providing choice with respect to both is likely to be beneficial.

Amelia and Emily

Emily was allocated the harder version of the audiobook, and noted that the length and difficulty level seemed about right, and the repetitive nature of the story (despite having the less repetitive version) was noticeable and advantageous. The audiobook was easy to administer, however engagement throughout the week was mixed. Amelia thought that the finger puppets (not utilised here due to logistics) might have had a positive impact on engagement levels.

Notes: First time they listened was after nursery, and E was tired.

Notes: Would sometimes fidget and walk out.

Notes: Played every day, but she wasn't always interested. But she liked this more as they went through the week.

As with other parents, Amelia was keen to point out that Emily's normal routine included plenty of traditional shared reading. Emily was also familiar with audiobooks, and at bedtime would often choose an audiobook to listen to after she has been read to by an adult. This did not happen every day however, and Amelia felt that Emily engaged more consistently with the audio format when it was over a longer time-period and there was a wider choice of books available for her to choose.

Engagement incorporates numerous facets that are not necessarily easily observed, even when done in a systematic way by a trained observer. Parents themselves commented that engagement was not always easy to gauge, particularly when listening took place within a context that the child could not easily escape from (e.g., the car or the bath), and pondered over the importance of this: did it matter as long as the children were exposed to the words from the audiobook? The debate over whether children learn from overheard speech (speech that is not directed at them and not using any of the qualities of child-directed speech such as appropriate language and tone) is not a new one (see Chapter 2), and whilst an audiobook playing in the background is not overheard speech in the strictest sense, if the child has no interest in listening, it is likely that they will become distracted by other activities, with listening becoming more passive, words less likely to be retained, and the audiobook akin to distant overheard speech. Where a child has chosen to listen, motivational theories (e.g., Self-Determination Theory, Ryan & Deci, 2000) suggest that they will be more likely to actively engage with the audiobook and retain the information heard.

Parental report of engagement here was not formally measured, and included anecdotal observations such as a desire to stay within earshot whilst the story was playing, pressing play independently of adult instruction, and the presence of extra-textual talk during and outside of listening time. This generally positive report of interest and engagement with the audio format is verified by existing research suggesting that some children find this format engaging and enjoyable. For example, Zientarksi and Pottorff (1994) found a group of secondary-school-aged struggling readers reacted with renewed interest and enthusiasm when presented with the audiobook format, although these children were notably older, and their renewed ability to access more difficult texts may have been partially responsible for their enjoyment: thus, more research is required that focuses on younger age-groups (Best, 2020). Importantly, the fact that measures of engagement are anecdotal and grounded in parental perceptions rather than systematic procedure is of little consequence here; ultimately, it is within the parents' power to decide on the type of home learning environment they provide for their child, and if they perceive a particular activity to be engaging for their child, then this is highly likely to influence their decision about whether or not to facilitate that activity on a regular basis.

As noted by Amelia, children's engagement with the audiobook was not a dichotomous thing and tended to fluctuate. There were many contextual factors (e.g., tiredness or sibling distractions) that can affect how well children engage from one listening event to the next (or often within the same listening event). This is a practically relevant theme because ascertaining how well the audiobook engaged the children is difficult to do over the short period of time available for this project (one week per participant dyad), and short intervention durations have been noted as important in previous studies (e.g., Noble et al., 2020; McGillion et al., 2017). Expecting full engagement each time

the audiobook is played is not realistic or reflective of how children function, and fluctuating engagement could still equate to a steady level of engagement over time.

Emily's version of the audiobook incorporated her own name in place of *Goldilocks*, and a place name with which she was familiar. Contrary to intention, Amelia felt that this inclusion was not necessarily helpful, possibly because Emily knew the story well already and this predictability had been interrupted by these changes. With respect to the narrator's voice, Amelia reported that Emily had recognised that it was a man's voice, but didn't see particularly interested beyond that. She felt that using different accents or possibly a child's voice could be an idea worth pursuing.

> *Notes: She was confused about E_____ being in the story Goldilocks – a story she was already familiar with.*

Participants' experiences of including elements of familiarity (name and place) have thus been somewhat inconclusive. Parent report of either verbal acknowledgement or a noticeable interruption in the activity of their child at the moment of hearing suggested that children did attend more to the familiar elements of the story, but parents were unclear whether this had led to greater engagement beyond the initial recognition, and, as noted above, Emily felt the inclusion had been confusing more than anything. Indeed, perhaps the main beneficiaries of introducing familiar aspects into a story would be the (hypothetical) children who tend to engage with books very rarely, but it is difficult to ascertain this without further investigation within this group.

Action Taken (3): Changing the Narrator

The input from Caroline and Amelia provided further invaluable feedback to that already provided by the earlier cycles. Particularly notable were conversations around Theme 3 (learning and engagement). Both parents felt that engagement was sometimes good, but tended to fluctuate and depended on the context - and this didn't necessarily mean that the child wasn't interested – as to expect 100% engagement with the same story every day for one week wasn't a practical expectation (and the short time span involved is a limitation of the current project). Provision of variety and choice within the resource was again discussed as a way of ensuring engagement over time. Engagement was also talked about in terms of the difficulties gauging it, and to what extent it mattered if children could learn to some extent from overheard speech, and in some contexts were listening in contexts from which they could not escape (e.g., car or bath). The researcher also reflected on the pivotal importance of parental perceptions (with respect to their child's engagement or anything else), in that it is they who make decisions about the home learning environment, and what resources will form part of that. As noted above, there were contrasting perspectives on the inclusion of the familiar name and place, with Caroline reporting positive feedback, and Amelia reporting a more confused response, linked to Emily's familiarity with the story and the family's existing established routines for shared reading (Theme 1). On the other hand, upon further discussion about the idea of incorporating familiarity as a way to identify with a story, both parents thought that exploring different types of narrating voice would be worth pursuing (although Amelia wasn't sure her own child would notice this particular aspect).

The elements of familiarity included for this round, in the form of the child's name and place of residence, were simple and suitable for the scale and logistics of the project. More comprehensive adaptations are possible and could have further-reaching positive influences in helping parents and

children from specific groups to identify and engage with stories in audiobook format or otherwise. Adhering to the approach that simple changes were easily actionable, thoughts and conversations returned to the voice of the narrator. Using a younger voice to narrate the story had been endorsed as an idea by several parents, and was something that could be easily tested in the present context. Whilst research has shown that babies prefer a familiar voice (DeCasper & Fifer, 1986), there does not appear to be any clear evidence on whether children attend more or less to unfamiliar children's voices as opposed to unfamiliar adults' voices. On the other hand, research does suggest that peers and siblings (thus likely to be closer in age) are known to play an important role in various realms of a child's development, including language (Yeomans-Maldonado et al., 2019).

Thus, the third and final change to be implemented was to re-record the audiobook using a child's voice. The intended outcome of this was to see whether the child's voice appeared to have a positive impact on engagement, as suggested by the significant role played by peers and siblings across a number of developmental spheres (Yeomans-Maldonado et al., 2019; Giletta et al., 2021). Not irrelevantly for a small-scale, time-constrained project, this change was also quick, cheap and uncomplicated to action in that it could be feasibly applied to the next cycle. This is in comparison to another possible adaptation - varying the narrating voice by accent – which would have involved both matching correct accents to the next participants, as well as finding a willing narrator with the correct matching accent. Furthermore, the fact that this batch of participants erred towards neutral, non-regional accents suggested that the use of a matching regional accent may not have had the same intended effect as for those with stronger regional accents.

Cycle 4: a Child Narrator

The final cycle of Action Research involved only one participant dyad – Lisa (mum) and Nathan (child) - who were allocated the harder version of the audio containing a greater diversity of verbs. This time no familiar elements were included, but the original script was narrated by an 11-year-old girl. The finger puppets were not provided for logistical reasons, as the participants had a specific time-slot for which they were available, and to arrange for a postage and pick up at an agreed location would have delayed and potentially reduced the time which they had available to take part.

Lisa and Nathan

Nathan listened to the audiobook every day for one week, and Lisa reported his initial engagement to be very good. She wondered whether this might have had something to do with the fact that the voice was coming from her phone, an object that was usually out of bounds for Nathan. Beyond the initial listening, engagement fluctuated somewhat, and wasn't always easy to ascertain. Lisa provided verbal encouragement, but the choice remained with Nathan as to whether he stayed listening, and at times he would talk whilst the audio was playing, or play with his brother. Lisa presumed that on these occasions he was completely disengaged from listening.

> *Notes: First time very interested – played on phone – perhaps the novelty of this as isn't normally allowed to play with phone, so isn't used to it. Wanted to press screen, etc.*
>
> *Notes: The second time he talked over it and wasn't so interested; went to play with his brother.*

As with some of the other child participants, Nathan was also familiar with the audio format, and had his own Yoto audio player. Lisa commented that the physical presence of the Yoto player, and the action required to insert the discs was an important feature of the player that Nathan liked, and helped to initiate his engagement with the story about to begin. Whilst the parent's phone (from which the audiobook was designed to play) was a physical entity, this was not a 'toy' item. This is an interesting point, and demonstrates a discordance between whether a particular aspect of the intervention is a benefit or a hindrance: playing the audiobook straight from a smartphone was intended to facilitate ease of play, and yet for parents who were opposed to their young child becoming familiar with smartphones, this unintentionally becomes a hindrance. Indeed, the provision of choice and independence intended here (in that children could start and stop the playing process themselves, in the same way that children could choose and insert the discs in the Yoto player) was likely frustrated in situations where children were not allowed physical access to their parent's phone, which is unfortunate as providing autonomy and decision-making power can be highly motivating for children in a learning and achievement setting (Self-Determination Theory, Ryan & Deci, 2000). Clearly working on the assumption (based on evidence of almost universal levels of internet access in the UK, Ofcom, 2023) that all parents are happy with convenient and technologically based learning resources is misguided to a certain extent, and will depend on parenting approaches and overall family context.

Lisa thought that using a child's voice to narrate the story was well-received by Nathan, although a level of uncertainty existed about the magnitude of its influence. This uncertainty could be extended to many aspects of the audiobook, in that opinions expressed tentatively on behalf of children cannot be said to be their own, and yet, due to the pivotal role they play in their child's upbringing, the parent's opinion is equally or indeed sometimes more important. Thus, if a parent is of the opinion that a child narrating an audiobook is beneficial, either because of the interest it may elicit in their child, or for any other reason, then this is likely to influence their decision regarding whether or not to acquire and play such an audiobook. Nonetheless, Lisa felt that the idea had potential, in the same way that using familiar or different local accents might be effective – essentially increasing the variety of what is out there, so that people (parents and children) can choose what suits them. For some families, accent or specifics about the content of the story will be important, and for others less so.

> *Notes: Depends on child and family. Perhaps providing a choice would be a good idea. Children have different likes/dislikes and families have different needs and routines.*

Again, Lisa was keen to emphasise that Nathan had an established shared reading routine, and was highly engaged with traditional reading books, often 'reading' on his own.

> *Notes: Hypothetically, L feels that audio could be good for some families who perhaps don't have such an established routine, or as a complement to families who already do lots of reading.*
>
> *Notes: …for her it definitely comes second best to traditional, interactive reading of a physical story book.*
>
> *Notes: L thinks this has something to do with the physical presence of a book.*

The aim of this research was to explore how effectively a vocabulary intervention could be delivered by parents in a simple audiobook format. This necessarily incorporated the discussing of potential challenges associated with using audiobooks, one of which presented itself in the guise of a

strong preference for traditional shared reading expressed by the majority of parents, and demonstrated clearly by Lisa, above. This was to some extent based on the perception of their child's preference, but also likely to be based upon their own idea of what constituted quality time spent together. It was not the intention to present audiobooks and physical books as mutually exclusive, however it is possible to see how this message became misconstrued along the way, with the primary focus of the project being audiobooks and their potential for positively influencing children's language. The consistent reiteration of this focus may have had the unintentional effect of pitting the two against each other, and parents were quick to point out the perceived disadvantages of audiobooks, for example a lack of both opportunity for interaction and physical presence.

Parents' unambiguous reaction to the (unintended) inference that audiobooks had the potential to become a superior alternative to shared reading with a physical book is interesting to explore further. Parents made clear and unsolicited assertions about their child's established reading routines, the presence of books in the house, and which format of reading was preferable for both themselves and their child (e.g., Sam in Cycle 1; Caroline in Cycle 3) - even when they had been more tentative about speaking on their behalf about other topics. Why might this be? Several reasons present themselves. Firstly, shared reading isn't just about the reading, but a shared bonding experience and physical closeness; a moment of calm amidst the tiredness and constant activity of parenting a young child (National Literacy Trust, 2020). Shared reading establishes familiar routines and shared enjoyment, and projects both backwards, in terms of nostalgic reminiscence, and forwards, in terms of passing on family practices and rituals to future generations. Traditional shared reading, therefore, is associated with so much more than just reading for both parents and children, and is thus a powerful tool that can be harnessed for continued engagement. Do audiobooks have the potential for similar powerful associations? The fact that a parent need not be physically close by for children to access audiobooks (indeed this can be interpreted as a benefit in certain contexts) probably removes some of the potential for associations with family bonding and communal enjoyment. Yet it is suggested that audiobooks can also facilitate enjoyment and engagement with a text (Best, 2020), and they certainly have the potential to be enjoyed communally in a similar way to other forms of audio. A good historical example of this is communal radio listening during the second world war, where people (albeit with limited choices available) frequently bonded and subsequently reminisced together over shared audio experiences.

Secondly, for many parents, shared reading is associated with both positive parenting and outcomes for children (Berkule et al., 2008), and parents are often aware of this. As previously noted, many of the discussions that took place during this Action Research project diverged into talking about hypothetical children and family contexts: particularly those who didn't have established shared reading routines, and whether and how these children may benefit from audiobooks. These discussions were extremely helpful, but possibly served to establish an 'us' and 'them' camp, with only the 'them' camp really requiring the audiobooks to assist in their language development. It was perhaps therefore important for parents, even on a subconscious level, to verify their allegiance with the 'established reading routine' group. Both a preference for shared reading and established routines for reading (as well as other activities) are important to consider (as per Theme 1), as they may have influenced both the child's engagement with the audiobook, as well as parental opinions – both of which are of interest for this Action Research project, although the direction of influence is not completely clear. For example, in Cycle 2 it was discussed that a child who has an established reading

routine is likely to be more capable of listening to and engaging with stories presented in a different format (audio) than a child who does not, probably due to well-developed listening and attention skills, and knowing what is expected of them. This observed engagement of a child by their parent is likely to positively influence their opinion of the audiobook, as their child appears to be listening and attending to the content. On the other hand, if a parent has a clear preference for traditional shared reading, and/or believes their child to have this same preference, and that audiobooks may interfere with this activity as described above, then this could have the opposite effect, in that it may negatively influence their opinions and their facilitation of the resource. It is important to emphasise that these suggestions are highly tentative, and in no way intended to abuse any narrative privilege, and yet they provide additional support for the need for further inquiry amongst families who do not have established reading routines, nor a strong preference for the traditional format.

Ultimately, traditional shared reading has been robustly associated with benefits on numerous different levels (Senechal & Lefevre, 2001); whether or not audiobooks can provide a similarly nurturing experience is less clear. To return to the point of earlier, any suggestion that audiobooks and traditional shared reading should be mutually exclusive was unintentional – the idea is first and foremost for audio to complement traditional shared reading, and in situations where the latter is not happening (only 58% of 3 to 4-year-olds are read to daily or nearly every day: David, 2020), audiobooks at least provide some exposure to language and narrative, as well as the potential for interaction outside of listening times. Despite the strong preferences for shared reading asserted, the idea of audiobooks complementing existing routines (their own, not a hypothetical family's) was embraced by some participants.

Summary of Cycle 4: Lisa and Nathan

Lisa's comments mirrored earlier feedback with respect to difficulties in gauging fluctuating engagement, as well as reiterating the importance of physical resources for more experiential learning and the importance of choice on every level (accents, format of resource, genre) to suit the individual preferences and needs of different families. There was a mixed response from Lisa regarding how influential having a child as narrator was: positive and well received overall, but unsure as to the magnitude of the influence. Using children to narrate audiobooks is relatively rare and unfamiliar to even those who are familiar with commercial audiobooks, and further research is likely required here. Consistent with the other parents/co-researchers, a clear preference for traditional shared reading was both expressly stated and implicitly evident, and the researcher reflected on why this might be the case, alongside reiterating the pivotal importance of parental opinions with respect to any part of a child's learning environment, including the resources which they decide to buy or acquire.

Finally, Lisa presented an alternative perspective regarding the use of mobile phones as a learning tool, which is somewhat at odds with the idea of harnessing technology (and specifically smartphones) to facilitate interventions for ease and convenience (Rowe et al., 2021). It should thus be remembered that despite the rising tide of mobile phone use and ownership, not all parents wish their young children to access this technology – although audiobooks played via alternative platforms, including the toy-like Yoto player, were presumably more acceptable. Nonetheless, Lisa related Nathan's initial engagement to this very point – access to something which he wasn't normally allowed to access. Again, this is somewhat different to the idea discussed by other participants/co-researchers above when relating engagement to routines with which they were already familiar (for example

engaging with the audio because they were familiar with Yoto players) – in that engagement or interest could also happen, particularly in the initial stage, when there was something incongruous or unfamiliar, which is in line with Hidi and Renninger's four phase model of interest development (Hidi & Renninger, 2006).

Conclusion

The aim of this Action Research project was to explore the feasibility of a simple parent-facilitated intervention in the form of a short audiobook, with the aim of increasing a child's exposure to verbs and improving their overall language acquisition (although the latter would not be tested here). This exploration was concerned with parental opinions about the format specifically: in terms of its feasibility as well as their child's perceived engagement. Such an intervention would then have the potential to address the vocabulary gap that exists between children with small and large vocabularies, although further research would be required with children across the vocabulary spectrum to see whether this was in fact the case. The reason for focusing on verbs, as noted throughout, is due to the finding in Phase One of this thesis that verbs seem to be particularly important in predicting overall vocabulary size, and could be the driving force behind additional word acquisition – therefore assisting those with smaller vocabularies to expand on the words they know and use. One of the key issues hoping to be addressed within this context was parents' and children's long-term motivation and engagement, and the ability of the intervention to 'nudge' them into new and sustainable habits.

Parent-child participant dyads were recruited two at a time and provided with one of two versions of a short audiobook resource, one with a narrower selection of repeated verbs, and the other with a more challenging selection of non-repeated verbs. They were asked to play the audio resource to their child at least once per day over a one-week period, before providing their feedback via an informal conversation. Changes were then made before handing the resource over to the next two participant dyads, with notes from these conversations forming the data to be analysed and feeding into the accumulating data pool to which Thematic Analysis was iteratively applied, with themes reviewed each time new data was added. Patterns of feedback obtained from parents, and consideration of existing theory and research alongside the researcher's own reflective input (as required by AR) fed into the decision process for any changes made, as well as the narrative account above. As described at the outset of this chapter, the account was constructed via a three-layer process, intended to combine a mostly linear narrative description of the findings, an analysis and discussion incorporating literature and theoretical explanations, and an addressing of the quality criteria described in the previous chapter. The final themes (Existing Routines, Ease of Audio, and Engagement & Learning) have been interwoven throughout the narrative. The findings, discussed in detail above, are summarised below, with associated limitations and implications for further research noted alongside.

Summary of Findings with Associated Limitations

Under the banner of Theme 3 (Engagement & Learning), a number of topics were considered, including parental perceptions of overall engagement and learning, and different aspects which may influence these, for example length, difficulty level, the addition of physical props, elements of familiarity and the importance of choice. Children's engagement with the audiobook, as observed by

the parent in an anecdotal, informal way, was generally good, but this was not the whole story. Engagement often fluctuated and was not always easy to gauge. Subjectivity, whilst embraced as part of Action Research methodology, should also be acknowledged here: one parent's 'fluctuating engagement' may have been another parent's 'good engagement,' and this is an important consideration if decisions are going to be made and action taken based on parental report of engagement. A limitation of the current research is that it relied solely on parental input with no objective element (whilst acknowledging the pivotal importance of parental perception in influencing decisions about their own children). Future research should seek to combine an objective element for triangulation purposes, for example observations of the child by a trained and objective observer. It should also allow a time-period of greater than one week per participant dyad to get a sense of engagement with audiobooks over a longer period of time, and beyond the period of novelty. Whether or not audiobooks could facilitate long-term positive habit changes towards providing an optimal language input was an important issue that this AR study hoped to address.

The audiobook was considered to be generally suitable in terms of its difficulty and length, although many parents noted that for audiobooks to be incorporated into real-world settings, an element of choice would be required. The use of repetition within the story appeared to be beneficial, and therefore care should be taken to balance the benefits of increasing verb diversity against the benefits of repetition (greater verb diversity can result in less overall repetition). A balance should also be struck between physical and non-physical aspects of the audiobook. Introducing some sort of associated physical presence, for example toys, storyboards or a physical audio 'player' (in this case finger puppets) was positive in that it seemed more likely that extra-textual talk and interactions would ensue, and focus the contingency back towards the audiobook. The finger puppets had the added benefit of being designed for action, so lent themselves towards re-enactment of story scenes and the emphasis of action verbs.

Both a parent and child being able to identify with a story and how it is delivered is important, although exactly how important probably depends on the individual circumstances of the family, and there was mixed feedback from the current sample, who in general acknowledged the importance of this issue, but did not think it overly relevant to their own context. Incorporating familiar content into the story (e.g., the child's name) is a simple way to facilitate identification with a story, but numerous potential adaptations exist (such as varying the accent of the narrator or adapting content more radically) that could be tailored to appeal to different groups. These more complex adaptations could have a greater influence than the simplistic ones used in this project, and should be explored in future research. Ultimately, the provision of choice to suit individual differences in various guises would be advantageous.

Existing routines, as outlined by Theme 1 (especially those around shared reading and access to books, but also general parenting approaches regarding choices and independence) were also a particular focus of the feedback discussions. All the participants in this study had established traditional shared reading routines and their children had plentiful access to books at home, although experience of audiobooks was mixed. Expectations of children's behaviour and parenting styles were not explicitly discussed, but authoritative styles appeared to be dominant, with parents seemingly warm, responsive and communicative, with high expectations of behaviour and independence. These existing routines are likely to be a factor in how well children engaged with the audiobook, in that

children who were used to engaging with shared reading and various other interactions with adults were also more likely to engage with the audiobook resource too. Further research needs to be done to see whether the findings here can be generalised to children who are not used to such experiences, or whose parents do not express a clear preference for traditional shared reading (although it should be noted again here that small-scale AR projects do not make claims of generalisability, but to contribute to knowledge).

In relation to Theme 2 (Ease of Audio), one of the benefits of audiobooks (beyond them being a potentially effective way for children to learn vocabulary), and part of the rationale for why they have been incorporated into this Action Research design - is their versatility. Audiobooks can be utilised in different ways, to different extents and for different reasons depending on a family's characteristics, preferences and routines. For example, in Chapter 7 it was noted that audiobooks may be particularly appealing to younger parents who have grown up with increased engagement with technology (Waite, 2018), and for busy parents who are required to constantly multi-task (Clark & Picton, 2018). On the other hand, parents who have strongly established reading routines (as was the case with this sample) are likely to use audio in a peripheral way, to complement existing routines, or possibly not at all. Audiobooks may also be particularly useful tools for parents who are struggling readers themselves, or have low motivation to read with their child, and therefore tend to avoid traditional shared reading. As previously noted, only 58% of 3 to 4-year-olds are read to daily or nearly every day according to data quoted by the National Literacy Trust (David, 2020), and this is even less for younger children (45% for 0 to 2-year-olds). Data collected by the BookTrust (2023) indicates that whilst 77% of families begin sharing books with their children before their first birthdays, this is only 23% of low-income families, despite 95% of parents acknowledging the importance of reading. These statistics, including the fact that 28% of parents reported not finding reading easy with their child, suggests that this group of families (those that don't read with their children), may be a particular beneficiary of audiobooks – especially if this encourages interaction between parent and child, but even, to some degree, if not.

Parents/co-researchers in the current project almost unanimously agreed audiobooks to be both accessible and convenient, although admittedly this does not necessarily translate into them being liked or used more by them, but simply adds to their 'nudging' potential. The use of technology (particularly a smartphone) to facilitate children's language development was also met with some reservation, both in terms of technological barriers, but also because the use of a smartphone to deliver the audiobook can be conceived as both a convenience and a hinderance - particularly where the phone is required for other tasks, or parents do not wish their child to handle or become fixated by their phone at such a young age. Audio not requiring the physical presence of a parent can also be interpreted in dichotomous ways: good for busy parents or for those who won't or can't, but bad for those who relish the bonding opportunity of time spent together. These discordances between whether a particular characteristic is a benefit or a hindrance depend on the specific family context. Individual differences between children, parents and family circumstances make it difficult to provide a one-size-fits-all resource (in terms of content and format), and is thus a limitation of this Action Research project. Future research on a larger scale should provide choices to accommodate these differences, both in terms of the platform the audiobook is provided through, but also in terms of any accompanying physical resources, and the range of texts available.

Of relevance to all three themes, parents want to be and be seen as good parents, but ideas about what is 'good' may differ depending on circumstances and background. Parents inhabit a unique position of importance with respect to their child's upbringing and language development, and it is they who make the decisions about what their children will do, read and listen to, hence their thoughts and opinions about audiobooks are vital. Having said that, the thoughts and feelings of the child are also important, and to some degree will influence the decisions their parents make. In this Action Research, children's perspectives were regularly considered through the eyes of their parents, whilst acknowledging their opinions were not always easy to determine, and sometimes the perspectives of parent and child became interwoven, for example concerning the preference for traditional shared reading. A limitation here was that the child's perspectives were not explicitly sought through direct questioning: future research should therefore attempt to consider both parent and child perspectives separately, given their importance.

Traditional shared reading is associated with positive emotions through shared interactive experiences. Audiobooks may have the potential to harness some of these positive associations, but ideally should be used alongside traditional shared reading and not as a replacement. Anecdotally, all parents in this sample reported regular reading routines with their children, and whilst no language testing took place, the reading routine alone suggested that these children will be likely to have good language outcomes in the future (Senechal & Lefevre, 2001). This is important, as whilst audiobooks can benefit and be enjoyed by everyone, perhaps their most important beneficiaries are the children of parents who (for a multitude of reasons) have not been able to establish optimal shared reading routines. As documented throughout the main discussion, participants often veered into hypothetical conversations about 'other' children and families such as these, and this initial feedback has been invaluable. Future research, however, should attempt to target those parents who have personal experience of the demographic factors associated with a higher risk of language delay or difficulties for their children.

Addressing Feldman's Quality Criteria

Action Research aims to take action, and to generate knowledge which is true and credible. To these ends, in Chapter 8 a number of quality criteria were outlined with a view to using said criteria as a framework and guide throughout the process, but also revisiting as part of the third layer of narrative construction to ensure that each criterion had been addressed. Each one will now be briefly considered, with associated limitations noted.

Reflexivity Principle. The principle of reflexivity requires AR to incorporate a critically reflective element. Specifically, this means questioning presumptions of knowledge and reality, making materials and methods transparent, and questioning the researcher's relationship with the overall research aims (incorporating a vocabulary intervention into an audio format). These aspects of the reflexivity principle were facilitated both independently by the researcher, as well as via conversations with critical friend (in the form of the researcher's PhD supervisor), and feedback and validation given by the wider supervision team. Questions of knowledge and reality, as well as a transparent description of the materials and procedure, were provided in Chapter 8. Consideration was also given to the researcher's own relationship with the research aims by positioning the researcher as a stakeholder who is responsible for the project and its outcomes and write up, and has values and experiences (particularly those relating to teaching children and of being a parent, and what has

'worked well') which will necessarily influence the whole AR process. Specific examples of where the reflexivity principle has come into play include utilising critical validation group feedback regarding making decisions and taking action (= a balance of feedback, research, researcher input and pragmatic factors), and critical friend conversations about having two versions of the audiobook to reflect how a future intervention study might look. The researcher's own relationship with the research aims (for example a belief based on previous experience of teaching and parenting that the audiobook would be a convenient and accessible solution for busy parents) were also questioned, particularly on occasions where contrary evidence appeared to be presented (for example Pippa experiencing barriers to facilitating the audiobook, and Lisa's aversion to using the smartphone as a vehicle for intervention and learning).

Dialectic Principle. The dialectic principle is concerned with the inclusion of multiple perspectives. In pursuance of this objective, this AR project firstly committed to presenting all data, including data which contravened the researcher's own view and/or the majority view. There are a number of examples of the latter in the narrative account provided above, including Pippa's minority view of the audio's convenience and ease of facilitation, and Amelia's assertion that inclusion of familiar names within the story was simply confusing. Secondly, it has also been suggested that the dialectic principle should incorporate a critical exchange of ideas (Feldman, 2007), and this was done by combining parental feedback, personal reflection and ideas collected during critical friend conversations. For example, the decision to incorporate elements of familiarity into the audiobook resource for Cycle 3 was a lengthy and complex consideration of parents' thoughts on the matter (important, but considered in a hypothetical way rather than relevant to their own children), the sort of changes which might work (accent, names, other features) related research and theory (e.g., interest theory), and pragmatic considerations (what changes would be feasible to achieve?) Thirdly, acknowledgement must be given to the narrative privilege held by the author, and due care was taken to avoid abusing such privilege by ensuring a balance between embracing subjectivity and preaching the moral high ground. This is particularly pertinent when dealing with information deemed to be sensitive or likely to be judged by others, in this case parenting behaviours. Fourthly, different ways must be sought to represent the same data, and to critique the researcher's own views. This was difficult to operationalise in some respects, but care was taken to discuss (via critical friend) and reflect upon all parental data provided, and use it to challenge views held. One example was the researcher's view that using different accents and voices for audio resources was likely to be an entirely positive action, in terms of having a wide appeal to different children and particularly those who struggle to identify with books of all formats. Caroline's assertion that accents should not be too strong initially felt in contravention of this view, and yet on further reflection, the concern that a narrating voice might not be easily understood by everyone was a worthy and pragmatic one.

Despite the evidence provided above of how the dialectic principle was operationalised within the current AR project, there are some potential limitations regarding how this was done. This project aimed to adopt a 'second-person practice action research' approach, which is focused upon 'engagement in collaborative work in co-inquiry and shared action with others on issues of mutual concern, through face-to-face dialogue, conversation and joint decisions and action' (Coghlan & Shani, 2021, p.468), with an emphasis on shared power and equality in contribution from all inquirers (McArdle, 2004, p.62). Did parent participants become co-inquirers as intended, and was there shared power and equality of contribution as indicated? Care was taken, as described above, to ensure

multiple perspectives were included and contributed to the decision-making process alongside other factors, and yet it was the researcher who was ultimately responsible for the project, in terms of making the final decisions, analysing the data and ultimately, for writing up – thus it would be remiss to assert that roles were either congruent or equal. As already noted, it is perhaps a limitation of the current project that data was not collected from a broader range of stakeholders beyond the parents and the researcher, including the children themselves, and others who have either specialised knowledge or interest in the topic. This, alongside following the quality criteria set forth by Feldman (2007) - which this AR project has done - might be the optimum way to keep the abuse of narrative privilege to a minimum (although avoiding it entirely might be impossible, given the subjective nature of AR).

Clear and Detailed Description of Data Collection. In their third quality requirement, Feldman (2007) suggests a clear and detailed description is needed for how and why the data has been collected. This description is provided in Chapter 8.

Construction of the Narrative. As noted in Chapter 8, there are many accepted ways to conceive and analyse data, as well as construct and present the narrative for AR projects. This project used an iterative process of Thematic Analysis (Braun & Clarke, 2006), with themes revisited each time new data was added from each set of participants, and then interwove these themes within three layers of narrative (as opposed to structuring the findings and discussion around them), so that the linear narrative of the change process could be retained. The first layer of narrative was a simple descriptive account of the feedback received from parents/co-researchers across multiple perspectives; a second layer of analysis and reflective discussion was then added; a final layer took the form of checking against the stated quality criteria, and adding details where required, for example making sure all data had been included, even where this may have contravened the researcher's/majority held views.

There were challenges presented by constructing the narrative, and the process was iterative to the very core. Thematic analysis lent itself as a suitable method for dealing with the data, particularly with respect to the slowly-expanding pool of data (due to the participants being recruited two at a time) and decisions made about changes being an ongoing process – however it was difficult to align the themes within a linear format, which was necessary to communicate the feedback and change process clearly. The decision was thus made to weave the themes throughout the layers of narrative. An alternative approach could have been to design the AR differently, recruit all the participant dyads at the same time, and ask them to listen to each version of the audio resource and provide feedback after each cycle, allowing more parental input to be fed into each decision regarding which changes should be made. On the other hand, pragmatic factors (such as the time available for the project, and the time available from parent participants) and the risk of parents and children become bored and desensitised to the resource, made this a less viable option. Ultimately, the 'messiness' and unpredictability of AR has been acknowledged in the previous chapter, and the decisions made regarding the analysis of data and construction of the narrative write up for the current project are not without limitation, but they have nonetheless incorporated feedback from critical friend and validation group, and stuck to the layered process described.

A further and related point, which could be considered a potential limitation, is in relation to the themes which emerged from the Thematic Analysis process, and how these fed into the changes made to the resource, an issue which was picked up on several times during critical friend discussions. All

themes were both rooted in the data and related to the research aim of exploring how effective the audio format could be, and yet the majority of changes made tended to be linked to Theme 3 (Engagement & Learning), and to a certain extent Theme 1 (Existing Routines). Theme 2 (Ease of Audio) was important for the purposes of the aims of this research, in that it was confirmatory to know that most the parents agreed that the audiobook was easy and convenient to facilitate, however the data did not transcend much beyond this agreement, nor translate well into concrete action that needed to be taken (in terms of making changes to the resource). On the other hand, if the scope of action was conceived in a broader way beyond making changes to the resource, then there are potential actions that could be linked to Theme 2 in terms of the broader encouragement of audiobook use in particular contexts (especially those requiring ease and convenience), for example on a school bus, or in a busy early years reading corner.

Inclusion of Explanation and Theory. The final quality indicator to mention (which has not already been incorporated in the discussion above) is the requirement to provide explanations, using established theory or otherwise, and critiquing these explanations. As part of the second narrative layer of this account, critical explanations have been woven into the narrative, including those linked to existing theory and research. For example, exploring the explanatory role of contingency and experiential learning (both important for vocabulary acquisition) with respect to the introduction of physical toys (the finger puppets) having the potential to improve a child's engagement with and learning from the audiobook.

Summary of Chapter 9

The aim of this Action Research was to explore how effectively a vocabulary intervention could be incorporated into a simple audiobook format by gathering parent feedback and using this in conjunction with theory, research and a subjective reflective contribution, to make and evaluate simple changes to the resource over a number of cycles. This chapter has provided a three-layered account of these findings, as well as revisiting the quality criteria presented in Chapter 8, and indicating where potential limitations exist, alongside ideas for future research. In summary, the audiobook has potential: it is accessible, convenient and often engaging, all of which are relevant in addressing the key issue of parents' and children's long-term motivation and the development of new and sustainable habits. The audiobook is also capable of eliciting extra-textual talk (especially with linked physical resources) and shared bonding experiences, yet it is unlikely to be an effective format across every context, and will depend on family and child characteristics: for example, incorporating elements of familiarity might be important, as will a family's existing routines. Parents play a pivotal role not only in what they do with their children, but also around the decisions they make about the home learning environment, including the resources which they choose. Whilst the aims of this Action Research phase of the project were not to test an intervention per se (in terms of its effect on children's vocabulary), but to explore how parents and children engage with the audio format – this could still perhaps be considered a limitation. Indeed, further consideration will be given to how an audiobook resource could be tested, couched in terms of suggestions for future research, in Chapter 10. This would also encompass the potential of audio for families who are at greater risk of experiencing language delay, or do not have established shared reading routines, as well as how choice and different ways of improving how children and families identify with stories might be incorporated. Ideally, the use of audiobooks should be alongside traditional reading materials.

Chapter 10: Bringing it All Together - a Proposal for Future Research

This current and final chapter will begin by summarising the key points from Phase One and Phase Two, before detailing a proposal of how future research might test the efficacy of an audiobook resource in terms of its ability to boost vocabulary. This proposal will be drawn from a combination of previous research alongside the findings from Phase One and Phase Two, and will attempt to address some of the limitations presented therein. Finally, the chapter will conclude this thesis by acknowledging the deficit perspective taken, considering how interventions can slot into our changing world and bridge the gap between academic knowledge everyday contexts, before a final note about going forward.

A Summary of Phase One and Phase Two

Phase One of this thesis analysed naturalistic video data of children's talk with the aim of finding out more about the nature of the vocabulary gap present between pre-schoolers with the smallest and largest vocabularies, important due to its robust associations with subsequent quality of life indicators (Gascoigne & Gross, 2017; Beitchman et al., 2001; Gilbert et al., 2018). Phase One also sought to pinpoint a more finely-grained picture of risk associated with language outcomes for children by exploring which demographic factors (such as parental age and household income) were predictive of vocabulary size. Two measures of vocabulary were employed to measure lexical diversity: VOCD (a measure of vocabulary diversity accounting for sample size) and NDW (the number of different words used standardised by time period of video recording). Verb usage emerged as a robust predictor of lexical diversity across both measures of lexical diversity, with rare word usage and word frequency also significant, but only for the NDW measure. There was a mixed profile of results with respect to which demographic factors were associated with children's vocabulary. VOCD was predicted by paternal age and paternal education, whereas NDW was predicted by gender, and no predictor was associated with both measures simultaneously. Furthermore, when additional analysis was carried out with similar demographic factors using the MCS dataset, this time almost all demographic variables included were predictive of vocabulary scores (no. of children, gender, income, paternal education, maternal age and maternal education), and thus the triangulation of results between the two datasets, intended to deepen understanding and provide an increased level of confidence in the findings, was unable to do so.. Possible explanations for the discrepancies in findings between both the two datasets, and the two measures of lexical diversity, have been discussed at length in Chapter 6, and it is clear that further research would be beneficial to unpick which factors are at play.

Phase Two looked to explore how the results from Phase One (specifically the most robust result: children who used a greater diversity of verbs were associated with having larger vocabularies overall) could be incorporated into an intervention which was aimed at boosting children's vocabulary, and ultimately narrowing the vocabulary gap. On reviewing the relevant intervention literature in Chapter 7, this intervention sought to embrace some of the strengths presented (e.g., utilising parents as facilitators, who are a cost-effective resource and well-placed to provide 1:1 support), whilst also circumventing some of the challenges presented, (e.g., ambiguity over whether fidelity requirements are sufficiently adhered to by parents, as well as the long term endurance of any positive effects). Within this context, a decision was made to explore a parent-facilitated audiobook format within an

Action Research design. Parental feedback was sought through four cycles of Action Research, and a narrative account of the findings combining layers of description, reflection and explanation was presented in Chapter 9. To summarise, the audiobook is accessible, convenient and often engaging, all of which may help nudge parents and children into sustainable habit change. The audiobook was also effective at eliciting extra-textual talk and parent-child interactions, particularly where linked physical resources were present, although the influence of including familiar elements into the story was unclear. How parents use audiobooks, and how influential they are at boosting children's vocabulary, will depend on many factors (including their existing routines around shared reading) – but overall, and in line with previous (albeit limited) research (e.g., Olson et al., 2024; Zientarksi & Pottorff, 1994), the conclusion of Phase Two indicated that the audiobook format was worthy of further exploration.

An Outline Proposal for Future Research

Phase Two of this project used an Action Research approach: it was exploratory in nature and concerned with resource development and potential rather than testing. Using a similar approach to that adopted in Olson et al. (2024), the following section will outline a research proposal detailing how an audiobook intervention might be tested in terms of its ability to boost vocabulary. The context and rationale for this proposal has been outlined throughout this thesis and is summarised above. In short, audiobooks appear to be both an effective and engaging format (as per Phase Two findings, and e.g., Woolfson, 2008), and show promise for influencing children's learning (Olson et al., 2024; Zientarksi & Pottorff, 1994) – however further research is required to see whether this extends to younger children, and applies to children from all backgrounds, particularly those without established shared reading routines. Furthermore, further research to confirm or otherwise whether exposure and learning of a broader range of verbs has a positive effect on overall vocabulary – as found in Phase One, would also be beneficial. Set against this context, therefore, the suggestion outlined below would aim to find out both whether listening to audio improved vocabulary overall (when compared to a control group), as well as whether there were any differences between listening to the version with greater verb diversity, as opposed to the one with less verb diversity. The sample would include children who experienced a wide variety of home learning environments and reading routines, and would continue for a duration aimed to maximise long-term impact.

Overview of Proposed Method: Sample, Materials and Procedure

To address the stated objectives, a RCT (Randomised Control Trial) design involving three conditions could be employed: audiobook with lesser verb diversity; audiobook with greater verb diversity; mindfulness control group. This is the same as the control group used in Olson's study, and whilst still an active control group, the activity has been deliberately chosen due to it being markedly different to the intervention activity, so as to avoid a possible moderation of the intervention effect as was the case in Noble et al. (2020). As per Olson et al. (2024), there would be three phases of the intervention: pre-test, intervention and post-test. As the study would be conducted remotely during all three phases, this would add to its potential reach and scalability (Olson et al., 2024). It would be expected that the audiobook conditions would see significantly greater improvements in children's language outcomes between pre- and post-test when compared to the control condition, with condition 2 (involving greater verb diversity) facilitating the greatest gains.

A sample of 3-year-old children and their families would be recruited. Whilst there would be no specific demographic indicators for recruitment, it would be important to ensure that a broad range of families were included within the sample, including a range of family sizes and home learning routines, particularly around reading. This recruitment could be done in partnership with other agencies, for example in collaboration with health visitors or early years childcare settings, thereby making use of data already collected to ensure a broad sample of families took part.

As in Phase Two, the audiobooks would be approximately 3 minutes long (which was judged to be about the right length in that phase), only this time a broader range of audiobook titles including both fiction and non-fiction genres would be provided, with additional titles released at different time points over the intervention period. Choice was an important component of Phase Two discussions, with the availability of only one audiobook title to listen to generally considered a limitation. Providing greater choice over a variety of genres would optimise the chances of engagement over a longer time period. The intervention phase would take place over a 30-week period. This was chosen as many researchers (e.g., McGillion et al., 2017; Noble et al., 2020) have noted that short interventions are not sufficient to influence vocabulary – indeed whilst Olson et al. (2024) found an effect on proximal measures of vocabulary after the 8-week audiobook intervention period, they did not find the same for standardised measures of vocabulary, thus suggesting a longer time period would be optimal. Two versions would be created of each audiobook title, with the same set of verbs used in each title as was used in Phase 2 (depending on the condition: lesser or greater verb diversity).

At pre-test, basic demographic information would be obtained from parents, including household size and information on their home learning environment and specifically reading routines (for example using the Home Learning Environment Index, Melhuish et al., 2001). Children would complete both standardised measures of vocabulary (with possible options including those described in Chapter 1, e.g., Renfrew Word Finding Vocabulary Test, British Picture Vocabulary Scale, and MacArthur-Bates Communicative Development Inventory) and proximal measures of vocabulary (testing knowledge of the included verbs, depending on which intervention condition they were part of). At post-test, the children would complete the same vocabulary measures. For the intervention phase, parents would be asked to facilitate the playing of the audiobook resources for 60 hours per week, over a period of 30 weeks (as justified above). The required listening hours per week is less than the 90 hours prescribed in Olson et al. (2024), but nonetheless accommodates shorter length audiobooks and younger children, and sets manageable expectations over a longer time period. Ideally, as in Olson et al. (2024), parents would be sent text reminders once or twice per day, and be offered some sort of compensation for their time, which would act as incentive by being linked to their levels of engagement with the project. For example, the audiobooks could be linked to an online platform which logged listening time, and when the recommended listening time had been reached each week (60 hours), this would release some sort of reward. This strategy might ensure some sort of fidelity in terms of encouraging the audiobook gets enough airtime, without dictating how this should be done. This same online platform would release new audiobook titles every fortnight to ensure continued choice and engagement. Compensating parents for their participation and engagement is not ideal in terms of nudging parents into sustainable habit changes, as it makes it difficult to ascertain their motives for engaging with the resource - however in terms of testing the audiobook for its ability to boost vocabulary, such compensation could incentivise engagement in terms of facilitating the required weekly listening times, thereby allowing the stated objectives to be addressed.

Limitations

As discussed in Chapters 8 and 9, audiobooks can be used in different ways. For example, some parents will utilise the audiobook in a similar way to a physical book, listening together with their child and interacting as they go along, with others leaving the child to listen alone whilst they get on with something else. Most parents will likely use a combination of both approaches, with styles of interaction varying between contexts (Cameron-Faulkner, 2020, p.174). These varying approaches and levels of parental input and interaction make it difficult to isolate the effects of the audiobook on vocabulary development, but, as with all intervention studies which require a degree of ecological validity, this is perhaps an accepted drawback, although future research should seek to isolate these contributions in more controlled conditions. Chapter 9 also outlined the potentially important role of using linked physical resources (the finger puppets), as well as the potential (although unconfirmed) importance of familiarising elements of the story to improve the way children identified with literacy materials. Future iterations of research should also bring these elements into the mix to explore whether vocabulary learning via audiobook can made even more effective by their introduction, and if so, whether this would be true across all children.

Acknowledging the Deficit Perspective

In Chapter 1, mention was made of the issues associated with employing deficit perspectives of children's development, by way of conceiving smaller vocabularies as a 'problem' that needs to be solved – as the current thesis has done. An alternative perspective would be to accept children's differences, and instead problematise our standardised, test-based education approach, and our system for selecting people for the right careers. These two opposing perspectives are difficult to reconcile, and striking a balance between the belief that children are best served by acquiring a diverse vocabulary, but without enforcing particular social and parenting ideologies for how to go about achieving this, is a sensitive and difficult issue, and deviates into epistemological realms on what can be considered truth and knowledge, and who should define what 'good,' 'optimal' and 'successful' look like. This is an important issue to acknowledge here - but nonetheless somewhat beyond the scope of this thesis.

Ultimately, whether agreeable or not, the system within which we all currently exist means that a child who has a delayed vocabulary upon entering school is likely to experience ongoing negative repercussions throughout life – or certainly what many people would judge to be negative, such as poorer job, health and life expectancy outcomes (Gascoigne & Gross, 2017; Gilbert et al., 2018). Surely what is required here is balance – approaches to teaching and intervening before school which appreciate people's diversity and backgrounds, celebrating and encouraging whatever input they are able to provide, whilst also making tentative suggestions about how to adapt input for optimal learning. Such suggestions could be made directly, using respectful and non-condescending language, or perhaps better still, through subtle adaptations to the environment to influence choices and nudge parents and children into effortless contexts which facilitate language learning, and thus enable the child to achieve the best they can within the system.

Bridging the Gap Between Academia and Practice

The pragmatic approach used throughout this thesis has been signposted at various stages (see Chapters 4 and 8). Using a pragmatic paradigm allowed for different yet aligned methodology across Phase One (corpus linguistics: quantitative analysis) and Phase Two (Action Research: qualitative analysis). Pragmatism facilitates a resolute focus on the issue to be addressed over and above any other considerations, and as such a flexibility in decision making which aligns with this focus (e.g., the decision to define verbs in a broad and practically relevant way, and the ability to incorporate a broad range of input - from parents, the researcher, existing theory and evidence - into the decision-making process for Phase Two). A single-minded focus on the real-world issue to be addressed also necessarily concerns itself with bridging the gap between academic research and practice.

For most children and families, intricate knowledge about the stages of children's language acquisition (e.g., how young children segment sounds from the speech stream, or how using the exaggerated tones of motherese and talking contingently can facilitate language learning) is not necessary. This is because most children in most families will be provided with enough verbal input for them to acquire language (and specifically vocabulary) with little effort, and without consideration of the underlying process. For those families with children who do not acquire language quite so effortlessly, imparting knowledge intended to support language development is only part of the equation: if families choose not to incorporate this knowledge into their way of life, or too many barriers exist for them to do so, then the knowledge becomes futile, and thus illustrates the difficulties of translating academic knowledge into useful practice in everyday contexts and relating to individual family circumstances. Results from Phase One of this thesis revealed that verbs are important facilitators of children's language acquisition – but how can this knowledge be utilised in the optimal way for the benefit of children with smaller vocabularies?

In response to this question, and to try and circumvent some of the barriers which some families experience in trying to provide an optimal verbal environment for their children, Phase Two has argued for the nudge. Nudge theory endorses subtle changes to the environment to influence people's choices, nudging them into make the 'best' choice, which is also the most effortless. The findings in Phase Two indicated that ease of facilitation was an important consideration for parents delivering language interventions in the pre-school years – and this might be especially true for those families experiencing multiple barriers, although further research is required across a broader range of family backgrounds to confirm this suggestion. By making subtle changes to the intervention approach (in this case, using an audiobook), elements of quality verbal input could be provided with limited effort.

Our Changing World

There is robust evidence for the current existence of a vocabulary gap (Fernald et al., 2013; Huttenlocher et al., 2010) but it is unclear whether the gap has widened or narrowed over time. Children's home learning environments are dynamic and liable to change with the times, but it is similarly unclear whether these changes are advantageous for children or not. Technological advancements (for example widespread smartphone ownership) may serve as additional distractions away from shared time between parents and children, or they may just be different ones to those experienced by earlier generations. The pros and cons of these societal changes seem somewhat of a moot point, however, as this is the world in which we exist, and it therefore seems pertinent to embrace - or at least harness – these changes (Rowe et al., 2021). This could be done by incorporating

language learning opportunities within the technology itself, and designing interventions (such as the audiobook) with busy working parents in mind. Innovative and sensitive solutions should also be sought to support the families who experience multiple barriers in facilitating their children's language development. Denmark, for example, has attempted to harness the strengths of the evolving demographics of their population (in terms of an increasing number of elderly or retired citizens), illustrated by the upward trend in the number of older adults volunteering across a range of settings (Amilon & Larsen, 2021). One example of this is the 'School Aunts and Uncles Program' (Tindbaek, 2019), where older volunteers have been utilised as a valuable resource in early years settings and schools, effectively increasing the amount of 1:1 time children receive with an adult (often via shared reading) – an arrangement which is likely to be reciprocally beneficial for both parties.

A Final Word: Going Forward

There are numerous ways that future research could build upon the outcomes of this thesis, with one example detailed earlier in this chapter in the outline proposal for future research. Alternative intervention approaches could also be explored by future research, for example making use of the natural link between action words and activity to facilitate verb learning, not dissimilar to other interventions which incorporate movement such as the Movement and Storytelling Intervention described in Chapter 7 (Duncan et al., 2019). Exploring a range of different types of language interventions is important, as it was clear from Phase Two of this thesis that despite many commonalities, parents also have different preferences, routines, knowledge and expectations when it comes to children's learning and upbringing (and this would presumably be even more evident in a larger, more diverse sample).

It is thus logical that different language interventions are needed to suit different families, necessitating, for example, a variety of formats (including audio), and choice within the same format (e.g., regarding story genre and familiarity of characters) to induce engagement and identification with the resource. An intervention which has worked for one child, and presents all the hallmarks of success, may inexplicably not work for the next child (one size doesn't always fit all). For many children and families, especially those for whom more traditional interventions have been ineffective, the nudge has real potential as a vehicle of intervention and facilitator of behavioural change which could benefit children's language outcomes. This could be via an audiobook, as demonstrated here, or, in line with the National Literacy Trust's Literacy by Stealth idea, any other way of harnessing culture and incorporating language and literacy into the things that children do anyway – including comics, sport and computer games. Future research should make exploring these avenues a priority.

References

Abbot-Smith, K., Imai, M., Durrant, S., & Nurmsoo, E. (2017). The role of timing and prototypical causality on how preschoolers fast-map novel verb meanings. *First Language, 37*(2), 186-204. https://doi.org/10.1177/0142723716679800

Adams, T. E. (2008). A review of narrative ethics. *Qualitative Inquiry, 14*(2), 175-194. https://doi.org/10.1177/1077800407304417

Adamson, L. B., Kaiser, A. P., Tamis-LaMonda, C. S., Owen, M. T., & Dimitrova, N. (2020). The developmental landscape of early parent-focused language intervention. *Early Childhood Research Quarterly, 50*, 59-67. https://doi.org/10.1016/j.ecresq.2018.11.005

Adlof, S. M., Catts, H. W., & Lee, J. (2010). Kindergarten predictors of second versus eighth grade reading comprehension impairments. *Journal of Learning Disabilities, 43*(4), 332-345. https://doi.org/10.1177/0022219410369067

Agdam, S. J., & Sadeghi, K. (2014). Two Formats of Word Association Tasks: A Study of Depth of Word Knowledge. *English Language Teaching, 7*(10), 1-12. https://doi:10.5539/elt.v7n10p1

Akemoglu. Y., Muharib, R., & Meadan, H. (2020). A Systematic and Quality Review of Parent-Implemented Language and Communication Interventions Conducted via Telepractice. *Journal of Behavioral Education, 29,* 282–316. https://doi.org/10.1007/s10864-019-09356-3

Akhtar, N., Dunham, F., & Dunham, P. J. (1991). Directive interactions and early vocabulary development: The role of joint attentional focus. *Journal of Child Language, 18*(1), 41-49. https://doi.org/10.1017/S0305000900013283

Akhtar, N., & Tomasello, M. (1996). Two-year-olds learn words for absent objects and actions. *British Journal of Developmental Psychology, 14*(1), 79-93. https://doi.org/10.1111/j.2044-835X.1996.tb00695.x

Altvater-Mackensen, N., & Mani, N. (2013). Word-form familiarity bootstraps infant speech segmentation. *Developmental Science, 16*(6), 980-990. https://doi.org/10.1111/desc.12071

Ambridge, B., & Lieven, E. V. (2011). *Child language acquisition: Contrasting theoretical approaches*. Cambridge University Press. https://doi-org.ntu.idm.oclc.org/10.1017/CBO9780511975073

Ambridge, B., & Lieven, E. V. (2015) A Constructivist Account of Child Language Acquisition. In B. MacWhinney, & W. O'Grady (Eds.), *The Handbook of Language Emergence* (pp.478-510). Wiley Online Books. https://doi.org/10.1002/9781118346136.ch22

Ambridge, B., & Rowland, C. F. (2013). Experimental methods in studying child language acquisition. *Cognitive Science, 4*(2), 149-168. https://doi.org/10.1002/wcs.1215

American Speech-Language-Hearing Association. (n.d.). https://www.asha.org/public/developmental-milestones/communication-milestones-19-to-24-months/

Amilon, A., & Larsen, M. R. (2021). Volunteer work among older adults in Denmark, 1997–2017: What can explain the continuous upward trend? *European Journal of Ageing, 18*, 17-28. https://doi.org/10.1007/s10433-020-00571-w

Anderson, L. A., & Linden, J. F. (2016). Mind the gap: two dissociable mechanisms of temporal processing in the auditory system. *Journal of Neuroscience, 36*(6), 1977-1995. https://doi.org/10.1523/JNEUROSCI.1652-15.2016

Acharya, K. P., Budhathoki, C. B., Bjønness, B., & Devkota, B. (2020). School gardening activities as contextual scaffolding for learning science: participatory action research in a community school in Nepal. *Educational Action Research, 30*(3), 462–479. https://doi.org/10.1080/09650792.2020.1850494

Aslin, R. N., Woodward, J. Z., LaMendola, N. P., & Bever, T. G. (1996). Models of word segmentation in fluent maternal speech to infants. In K. Demuth & J.L. Morgan (Eds.), *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition* (pp. 117-134). Psychology Press.

Baguley, T. (2009). Standardized or simple effect size: What should be reported? *British Journal of Psychology, 100*(3), 603-617. https://doi.org/10.1348/000712608X377117

Baker, L., Mackler, K., Sonnenschein, S., & Serpell, R. (2001). Parents' interactions with their first-grade children during storybook reading and relations with subsequent home reading activity and reading

achievement. *Journal of School Psychology, 39*(5), 415-438. https://doi.org/10.1016/S0022-4405(01)00082-6

Baldwin, D. (1995). Understanding the link Between Joint Attention and Language. In C. Moore & P. Dunham (Eds.), *Joint Attention, Its Origin and Role in Development* (pp. 131-158). Lawrence Erlbaum Associates, Inc.

Baldwin, D. A., Markman, E. M., Bill, B., Desjardins, R. N., Irwin, J. M., & Tidball, G. (1996). Infants' reliance on a social criterion for establishing word-object relations. *Child Development, 67*(6), 3135-3153. https://doi.org/10.1111/j.1467-8624.1996.tb01906.x

Ballarino, G., Bernardi, F., Requena, M., & Schadee, H. (2009). Persistent inequalities? Expansion of education and class inequality in Italy and Spain. *European Sociological Review, 25*(1), 123-138. https://doi.org/10.1093/esr/jcn031

Balota, D. A., Yap, M. J., Hutchison, K. A., Cortese, M. J., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English lexicon project. *Behavior Research Methods, 39*, 445-459. https://doi.org/10.3758/BF03193014

Banks, D. (2016, April 4). What is brain plasticity and why is it so important. *The Conversation.* https://theconversation.com/what-is-brain-plasticity-and-why-is-it-so-important-55967

Banks, S., & Brydon-Miller, M. (2019). Ethics in participatory research. In S. Banks, & M. Brydon-Miller (Eds.), *Ethics in participatory research for health and social well-being : cases and commentaries.* (pp. 1-30). Routledge.

Barbosa, C., Vasquez, S., Parada, M. A., Gonzalez, J. C. V., Jackson, C., Yanez, N. D., Gelaye, B., & Fitzpatrick, A. L. (2009). The relationship of bottle feeding and other sucking behaviors with speech disorder in Patagonian preschoolers. *BMC Pediatrics, 9*, 1-8. https://doi.org/10.1186/1471-2431-9-66

Barbu, S., Nardy, A., Chevrot, J-P., Guellai, B., Glas, L. Juhel, J., & Lemasson, A. (2015). Sex Differences in Language Across Early Childhood: Family Socioeconomic Status does not Impact Boys and Girls Equally. *Frontiers in Psychology, 6*(1871), 1-10. https://doi.org/10.3389/fpsyg.2015.01874

Bates, E., Bretherton, I., & Snyder, L. S. (1988). *From first words to grammar: Individual differences and dissociable mechanisms.* Cambridge University Press.

Bates, E., Dale, P. S., & Thal, D. (1995) Individual differences and their implications for theories of language development. In P. Fletcher, & B. MacWhinney (Eds.), *The Handbook of Child Language* (pp.95-151). Blackwell Publishing Ltd. https://doi.org/10.1111/b.9780631203124.1996.x

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., Reilly, J., & Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language, 21*(1), 85-123. https://doi:10.1017/S0305000900008680

Beals, E. D., & Tabors, O. P. (1993, March 25-28). Arboretum, Bureaucratic, and Carbohydrates: Preschoolers' Exposure to Rare Vocabulary at Home [Conference session]. Biennial Meeting of the Society for Research in Child Development 1993, New Orleans, United States. https://files.eric.ed.gov/fulltext/ED356057.pdf

Beaman, P. C., Hanczakowski, M., & Jones, D. M. (2014). The effects of distraction on metacognition and metacognition on distraction: Evidence from recognition memory. *Frontiers in Psychology, 5*(439), 1-13. https://doi.org/10.3389/fpsyg.2014.00439

Becker, J. A. (1994). Pragmatic Socialization: Parental Input to Preschoolers. *Discourse Processes, 17*(1), 131-148. https://doi.org/10.1080/01638539409544862

Becker, B. (2011). Social disparities in children's vocabulary in early childhood. Does pre-school education help to close the gap. *The British Journal of Sociology, 62*(1), 69-88. https://doi.org/10.1111/j.1468-4446.2010.01345.x

Behme, C., & Deacon, S. H. (2008). Language learning in infancy: Does the empirical evidence support a domain specific language acquisition device? *Philosophical Psychology, 21*(5), 641-671. https://doi.org/10.1080/09515080802412321

Behrens, H. (2021). Constructivist approaches to first language acquisition. *Journal of Child Language, 48*(5), 959-983. https://doi.org/10.1017/S0305000921000556

Beitchman, J. H., Wilson, B., Johnson, C. J., Atkinson, L., Young, A., Adlaf, E., Escobar, M., & Douglas, L. (2001). Fourteen-Year Follow-up of Speech/Language-Impaired and Control Children: Psychiatric Outcome. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(1), 75-82. https://doi.org/10.1097/00004583-200101000-00019

Bellugi, U., Lichtenberger, L., Jones, W., Lai, Z., & St. George, M. (2000). I. The neurocognitive profile of Williams Syndrome: a complex pattern of strengths and weaknesses. *Journal of Cognitive Neuroscience, 12*(1, Suppl.), 7-29. https://doi.org/10.1162/089892900561959

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language, 6*(2), 183-200. https://doi:10.1017/S0305000900002245

Bercow, J. (2008*). The Bercow Report: A review of services for children and young people (0-19) with speech, language and communication needs.* Department for Children Schools and Families. https://dera.ioe.ac.uk/id/eprint/8405/7/7771-dcsf-bercow_Redacted.pdf

Bercow, J. (2018). *Bercow: Ten Years On: An independent review of provision for children and young people with speech, language and communication needs in England.* I CAN, Royal College of Speech & Language Therapists. http://www.bercow10yearson.com/wp-content/uploads/2018/03/337644-ICAN-Bercow-Report-WEB.pdf

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences, 109*(9), 3253-3258. https://doi.org/10.1073/pnas.1113380109

Berkule, S. B., Dreyer, B. P., Klass, P. E., Huberman, H. S., Yin, H. S., & Mendelsohn, A. L. (2008). Mothers' expectations for shared reading after delivery: implications for reading activities at 6 months. *Ambulatory Pediatrics, 8*(3), 169-174. https://doi.org/10.1016/j.ambp.2008.01.002

Best, E. (2020) *Audiobooks and Literacy: A Rapid Review of the Literature.* National Literacy Trust. https://files.eric.ed.gov/fulltext/ED607775.pdf

Best, E. (2021). *Playful Storytelling: The Role of Interactive Audio in Building Children's Literacy Skills and Engagement.* National Literacy Trust. https://files.eric.ed.gov/fulltext/ED614755.pdf

Bierman, K. L., Torres, M. M., Domitrovich, C. E., Welsh, J. A., & Gest, S. D. (2009). Behavioral and cognitive readiness for school: Cross-domain associations for children attending Head Start. *Social Development, 18*(2), 305-323. https://doi.org/10.1111/j.1467-9507.2008.00490.x

Bishop, D. V. (2017). Why is it so hard to reach agreement on terminology? The case of developmental language disorder (DLD). *International Journal of Language & Communication Disorders, 52*(6), 671-680. https://doi.org/10.1111/1460-6984.12335

Blackwell, A. A. (2005). Acquiring the English adjective lexicon: relationships with input properties and adjectival semantic typology. *Journal of Child Language, 32*(3), 535–562. https://doi:10.1017/S0305000905006938

Blake, J. (1981). Family size and the quality of children. *Demography, 18*, 421-442. https://doi.org/10.2307/2060941

Blank, M. Gessner, M., & Esposito, A. (1979). Language without communication: a case study. *Journal of Child Language, 6*(2), 329–352. https://doi:10.1017/S0305000900002336

Blewitt, P., & Langan, R. (2016). Learning words during shared book reading: The role of extratextual talk designed to increase child engagement. *Journal of Experimental Child Psychology, 150*, 404-410. https://doi.org/10.1016/j.jecp.2016.06.009

Bloom, L. (1974). Talking, understanding, and thinking: developmental relationship between receptive and expressive language. In R. Schiefelbusch, & L. Lloyd (Eds.), *Language perspectives: acquisition, retardation and intervention* (pp.285-311). University Park Press.

Bloom, L. (1993). *Language development from two to three*. Cambridge University Press.

Bold, C. (2012). *Using Narrative in Research.* Sage Publications Ltd. http://digital.casalini.it/9781446254264

BookTrust (2023). *Children's reading habits in the early years.* The Book Trust. https://www.booktrust.org.uk/globalassets/resources/research/family-survey-briefing-1-reading-in-the-early-years-2023-1.pdf

Booth, A. E., & Waxman, S. R. (2009). A horse of a different color: Specifying with precision infants' mappings of novel nouns and adjectives. *Child Development, 80*(1), 15-22. https://doi.org/10.1111/j.1467-8624.2008.01242.x

Bornstein, M. H., Haynes, M. O., & Painter, K. M. (1998). Sources of child vocabulary competence: A multivariate model. *Journal of Child Language, 25*(2), 367-393. https://doi:10.1017/S0305000998003456

Bortfeld, H., Morgan, J. L., Golinkoff, R. M., & Rathbun, K. (2005). Mommy and me: Familiar names help launch babies into speech-stream segmentation. *Psychological Science, 16*(4), 298-304. https://doi.org/10.1111/j.0956-7976.2005.01531.x

Boshart, C. A. (2001). *The pacifier: Making the decision*. Speech Dynamics.

Boulton, J. & Preiser, R. (2021). Action Research. In R. Biggs, A. de Vos, R. Preiser, H. Clements, K. Maciejewski, & M. Schlüter (Eds.), *The Routledge Handbook of Research Methods for Social Ecological Systems* (pp. 217-229). Routledge. https://doi.org/10.4324/9781003021339

Bradbury, B., Corak, M., Waldfogel, J., & Washbrook, E. (2015). *Too many children left behind: The US achievement gap in comparative perspective*. Russell Sage Foundation.

Bradbury-Huang, H. (2015). *The Sage Handbook of Action Research*. (3rd ed.). Sage Publications Ltd. http://digital.casalini.it/9781473927247

Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. C. (2015, July 22-25). *Developmental Changes in the Relationship Between Grammar and the Lexicon.* [Conference presentation]. 37th Annual Meeting of the Cognitive Science Society, California, United States. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=ebccdecb304fc5a6c5b455ac352cb5c8820b02f3

Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. (2016 August 10-13). From uh-oh to tomorrow: Predicting age of acquisition for early words across languages. [Conference presentation]. 37th Annual Meeting of the Cognitive Science Society, California, United States. https://callab.github.io/publication/braginsky-2016-from/braginsky-2016-from.pdf

Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. C. (2019). Consistency and variability in children's word learning across languages. *Open Mind, 3*, 52-67. https://doi.org/10.1162/opmi_a_00026

Brand, J. E. (2015). The Far-Reaching Impact of Job Loss and Unemployment. *Annual Review of Sociology, 41,* 359-375. https://doi.org/10.1146/annurev-soc-071913-043237

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health, 11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Braunwald, S. R. (1995). Differences in the Acquisition of Early Verbs: Evidence from Diary Data from Sisters. In M. Tomasello & W.E. Merriman (Eds.), *Beyond Names for Things: Young Children's Acquisition of Verbs* (pp. 81-111). Lawrence Erlbaum Associates, Inc.

Breen, R., Luijkx, R., Müller, W., & Pollak, R. (2009). Nonpersistent inequality in educational attainment: Evidence from eight European countries. *American Journal of Sociology, 114*(5), 1475-1521. https://doi.org/10.1086/595951

Brent, M. R., & Siskind, J. M. (2001). The role of exposure to isolated words in early vocabulary development. *Cognition, 81*(2), B33-B44. https://doi.org/10.1016/S0010-0277(01)00122-6

Brock, J. (2007). Language abilities in Williams syndrome: A critical review. *Development and Psychopathology, 19*(1), 97-127. https://doi.org/10.1017/S095457940707006X

Brown, P., & Gaskins, S. (2014). Language acquisition and language socialization. In M. Dingemanse, S. Floyd, N. J. Enfield, P. Kockelman, & J. Sidnell (Eds.), *Cambridge Handbook of Linguistic Anthropology* (pp. 187-226). Cambridge University Press.

Bryant, J. B. (2009) Pragmatic development. In E. L. Bavin (ed.), *The Cambridge Handbook of Child Language*. (pp. 339-354). Cambridge University Press.

Brydon-Miller, M., Greenwood, D., & Maguire, P. (2003). Why action research? *Action Research, 1*(1), 9-28. https://doi.org/10.1177/14767503030011

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods, 46*, 904-911. https://doi.org/10.3758/s13428-013-0403-5

Burgess, A., & Goldman, R. (2022). *Bringing Baby Home: UK fathers in the first year after the birth.* Fatherhood Institute. http://www.fatherhoodinstitute.org/wp-content/uploads/2022/06/Bringing-Baby-Home-Main-Report.pdf

Burgess, J. (2006). Participatory action research: First-person perspectives of a graduate student. *Action Research, 4*(4), 419-437. https://doi.org/10.1177/1476750306070104

Burgoyne, K., Gardner, R., Whiteley, H., Snowling, M. J., & Hulme, C. (2018). Evaluation of a parent-delivered early language enrichment programme: evidence from a randomised controlled trial. *The Journal of Child Psychology and Psychiatry, 59*(5), 545-555. https://doi.org/10.1111/jcpp.12819

Burnett, C., Merchant, G., & Neumann, M. M. (2020). Closing the gap? Overcoming limitations in sociomaterial accounts of early literacy. *Journal of Early Childhood Literacy, 20*(1), 111-133. https://doi.org/10.1177/1468798419896067

Burr, S., Harding, S., Wren, Y., & Deave, T. (2021). The Relationship between feeding and non-nutritive sucking behaviours and speech sound development: a systematic review. *Folia Phoniatrica et Logopaedica, 73*(2), 75-88. https://doi.org/10.1159/000505266

Bybee, J. L., & Beckner, C. (2012). Usage-Based Theory. In B. Heine, & H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis* (pp. 827-856). Oxford Academic. https://doi.org/10.1093/oxfordhb/9780199544004.013.0032

Cabrera, N. J., Shannon, J. D., & Tamis-LeMonda, C. (2007). Fathers' influence on their children's cognitive and emotional development: From toddlers to pre-K. *Applied Development Science, 11*(4), 208-213. https://doi.org/10.1080/10888690701762100

Cameron-Faulkner, T., & Noble, C. (2013). A comparison of book text and Child Directed Speech. *First Language*, *33*(3), 268-279. https://doi.org/10.1177/01427237134876

Cameron-Faulkner, T., Lieven, E., & Tomasello M. (2003). A construction based analysis of child directed speech. *Cognitive Science, 27*(6), 843 – 873. https://doi.org/10.1207/s15516709cog2706_2

Cameron-Faulkner, T. (2020). The emergence of gesture during prelinguistic interaction. In C. F. Rowland, A. L. Theakston, B. Ambridge, & K. E. Twomey (Eds.), *Current Perspectives on Child Language Acquisition* (pp.173-188). John Benjamins Publishing Co.

Cameron-Faulkner, T., Theakston, A., Leiven, E. V., & Tomasello, M. (2015) The relationship between infant hold out and gives, and pointing. *Infancy, 20*(5), 576-586. https://doi.org/10.1111/infa.12085

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and reports on Child Language Development, 15,* 17-29.

Cartmill, E. A. (2016). Mind the gap: Assessing and addressing the word gap in early education. *Policy Insights from the Behavioral and Brain Sciences, 3*(2), 185-193. https://doi.org/10.1177/237273221665756

Cartmill, E. A., Armstrong III, B. F., Gleitman, L. R., Goldin-Meadow, S., Medina, T. N., & Trueswell, J. C. (2013). Quality of early parent input predicts child vocabulary 3 years later. *Proceedings of the National Academy of Sciences, 110*(28), 11278-11283. https://doi.org/10.1073/pnas.1309518110

Chall, J. S., & Dale, E. (1995). *Readability revisited and the new Dale-Chall readability formula.* Brookline Books

Chen, S. P., Horgan, S., Jones, J., Krauss, E., & Stuart, H. (2021). Involving student peer researchers for gender-informed health promotion: a community-based participatory action research. *Educational Action Research, 31*(3), 521–539. https://doi-org.ntu.idm.oclc.org/10.1080/09650792.2021.1970603

Chomsky, N. (1957). *Syntactic structures.* Mouton de Gruyter.

Chomsky, N. (1978). *Topics in the theory of generative grammar*. Mouton De Gruyter. https://doi.org/10.1515/9783110903843

Clark, C., & Picton, I. (2018). *Book ownership, literacy engagement and mental wellbeing.* National Literacy Trust. https://nlt.cdn.ngo/media/documents/National_Literacy_Trust_-_Book_ownership_report.pdf

Clark, C., & Picton, I. (2023). *Children and young people's book ownership in 2023: A 10-year retrospective.* National Literacy Trust. https://nlt.cdn.ngo/media/documents/Book_ownership_in_2023_k6ovIWY.pdf

Clark, E. V., & Hecht, B. F. (1983). Comprehension, production, and language acquisition. *Annual Review of Psychology, 34*, 325-349. https://doi.org/10.1146/annurev.ps.34.020183.001545

Cochet, H., & Byrne, R. W. (2016). Communication in the second and third year of life: Relationships between nonverbal social skills and language. *Infant Behavior and Development, 44*, 189-198. https://doi.org/10.1016/j.infbeh.2016.07.003

Coghlan, D., & Shani, A. B. (2021). Abductive Reasoning as the Integrating Mechanism between First- Second- and Third-Person Practice in Action Research. *Systematic Practice and Action Research, 34,* 463–474. https://doi.org/10.1007/s11213-020-09542-9

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Cole, A., Brown, A., Clark, C., & Picton, I. (2022). *Children and Young People's Reading Engagement in 2022: Continuing Insight into the Impact of the COVID-19 Pandemic on Reading.* National Literacy Trust. https://nlt.cdn.ngo/media/documents/Reading_trends_2022_-_Final.pdf

Coleman, L., & Cottell, J. (2019). *Childcare Survey 2019.* Coram Family and Childcare. https://www.familyandchildcaretrust.org/sites/default/files/Resource%20Library/Childcare%20Survey%202019_Coram%20Family%20and%20Childcare.pdf

Conica, M., Nixon, E., & Quigley, J. (2020). Fathers' but not mothers' repetition of children's utterances at age two is associated with child vocabulary at age four. *Journal of Experimental Child Psychology, 191*, 1-12. https://doi.org/10.1016/j.jecp.2019.104738

Connelly, R. (2013). *Millennium Cohort Study Data Note 2013/1: Interpreting test scores.* Centre for Longitudinal Studies. https://cls.ucl.ac.uk/wp-content/uploads/2017/07/MCS-data-note-20131-Test-Scores-Roxanne-Connelly.pdf

Crystal, D. (1987). *The Cambridge Encyclopaedia of Language.* Cambridge University Press.

Crystal, D. (2010). *A little book of language*. UNSW Press.

Curtiss, S. (2014). *Genie: A Psycholinguistic Study of a Modern-Day "Wild Child*." Academic Press, Inc.

Dadds, M., & Hart, S. (2001). *Doing practitioner research differently*. Routledge.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers, 28,* 125–127. https://doi.org/10.3758/BF03203646

Dale, P.S., & Kachergis, G. (2023). The CDI-III. In V. A. Marchman, P.S. Dale., & L. Fenson (Eds.), *MacArthur-Bates Communicative Development Inventories: Users Guide and Technical Manual* (3rd ed., pp. 125-134). Brookes Publishing.

Dale, P. D., & Patterson, J. L. (2017). Early Identification of Language Delay. In R. E. Tremblay, M. Boivin, R. Peters, & S. Rvachew (Eds.), *Encyclopedia of Early Language Development.* (pp.1-6). Centre of Excellence for Early Childhood Development and Strategic Knowledge Cluster on Early Child Development.

d'Apice, K., Latham, R. M., & von Stumm, S. (2019). A naturalistic home observational approach to children's language, cognition, and behavior. *Developmental Psychology, 55*(7), 1414-1427. https://doi.org/10.1037/dev0000733

David, A. (2020, Jan 2). Reading to children is so powerful, so simple and yet so misunderstood. *National Literacy Trust.* https://literacytrust.org.uk/blog/reading-children-so-powerful-so-simple-and-yet-so-misunderstood/

Davies, C., Hendry, A., Gibson, S. P., Gliga, T., McGillion, M., & Gonzalez-Gomez, N. (2021). Early childhood education and care (ECEC) during COVID-19 boosts growth in language and executive function. *Infant and Child Development, 30*(4), 1-15. https://doi.org/10.1002/icd.2241

Davis, J. M. (2007). Rethinking the architecture: An action researcher's resolution to writing and presenting their thesis. *Action Research, 5*(2), 181-198. https://doi.org/10.1177/1476750307077322

de la Rie S., van Steensel, R. C. M., van Gelderen, A. J. S., & Severiens, S. (2020). Level of abstraction in parent–child interactions: the role of activity type and socioeconomic status. Journal of Research in Reading, 43(1), 140-159. https://doi.org/10.1111/1467-9817.12294

de Wijk, R. A., Maaskant, A. J., Polet, I. A., Holthuysen, N. T., van Kleef, E., & Vingerhoeds, M. H. (2016). An in-store experiment on the effect of accessibility on sales of wholegrain and white bread in supermarkets. *PloS One, 11*(3), 1-8. https://doi.org/10.1371/journal.pone.0151915

deBoer, F. (2014). Evaluating the comparability of two measures of lexical diversity, *System, 47*, 139-145. https://doi.org/10.1016/j.system.2014.10.008.

DeCasper, A. J., & Fifer, W. P. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 208*(4448), 1174-1176. https://doit.org/10.1126/science.7375928

DeCasper, A. J., & Spence, M. J. (1986). Prenatal maternal speech influences newborns' perception of speech sounds. *Infant Behavior and Development, 9*(2), 133-150. https://doi.org/10.1016/0163-6383(86)90025-1

Denscombe, M. (2017). *The good research guide: For small-scale social research projects*. McGraw-Hill Education.

DePaolis, R. A., Keren-Portnoy, T., & Vihman, M. M. (2010, April 9-10). *A comparison of US and UK kids.* [Conference Presentation]. International Child Phonology conference, Memphis, United States.

Dewey, J. (1910). *How we Think.* D. C Heath & Co.

Dimova, S., Ilie, S., Brown, E. R., Broeks, M., Culora, A., & Sutherland, A. (2020). *The Nuffield Early Language Intervention: Evaluation Report.* Education Endowment Foundation. https://d2tic4wvo1iusb.cloudfront.net/production/documents/projects/Nuffield_Early_Language_Intervention.pdf?v=1717151576

Dockrell, J. E. (2019). Language Learning Challenges in the Early Years. In D. Whitebread, M. M. McClelland, K. Kumpulainen, G. Valeska (Eds.), *The SAGE Handbook of Developmental Psychology and Early Childhood Education* (pp. 435-452). Sage Publications Ltd. http://digital.casalini.it/9781526484567

Dockrell, J. E., Braisby, N., & Best, R. M. (2007). Children's acquisition of science terms: Simple exposure is insufficient. *Learning and Instruction, 17*(6), 577-594. https://doi.org/10.1016/j.learninstruc.2007.09.005

Dom, L. & Willockx, D. (2022). Allegories on creating opportunities for dialogue between childcare practitioners and parents during an action research in two out-of-school care centres. *Educational Action Research, 32*(2), 295-310. https://doi-org.ntu.idm.oclc.org/10.1080/09650792.2022.2096656

Dunatchik, A., Wishart, R., Cartagena-Farias, J., & Smith, N. (2018). *Regional differences in attainment in the early years.* National Centre for Social Research. https://assets.publishing.service.gov.uk/media/5b697c23e5274a14f45342d4/Regional_Early_Years_Attainment_Gap.pdf

Duncan, G. J., Lee, K. T., Rosales-Rueda, M., & Kalil, A. (2018). Maternal age and child development. *Demography, 55*(6), 2229-2255. https://doi.org/10.1007/s13524-018-0730-3

Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test* (4th ed.). Pearson Assessments.

Dunn L. M., Dunn D. M., Sewell J., Styles B., Brzyska B., Shamsan Y., et al. (2009). *The British Picture Vocabulary Scale* (3rd ed.). GL Assessment.

Durán, P., Malvern, D., Richards, B., & Chipere, N. (2004). Developmental trends in lexical diversity. *Applied Linguistics, 25*(2), 220-242. https://doi.org/10.1093/applin/25.2.220

Einboden, R., Maxwell, H., Campbell, C., Rickard, G., & Bramble, M. (2022). Improving the first-year student experience: a critical reflection on co-operative inquiry as the 'last loop' in an action research project. *Educational Action Research, 31*(5), 894–908. https://doi.org.ntu.idm.oclc.org/10.1080/09650792.2022.2034657

Elliott, C. D. (1996). *The British Ability Scales II* (2nd ed.). NFER-NELSON Publishing Company.

Elliott, C. D., Smith, P., & McCulloch, K. (1997). *British Ability Scales Second Edition (BAS II) Technical Manual.* London: Nelson.

Evans, J. S. B., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*(3), 223-241. https://doi.org/10.1177/1745691612460685

Ewert, B. (2020). Moving beyond the obsession with nudging individual behaviour: Towards a broader understanding of Behavioural Public Policy. *Public Policy and Administration, 35*(3), 337-360. https://doi.org/10.1177/0952076719889090

Eyre, C. (2019, September 23). Children's audiobooks double in popularity. *The Bookseller.* https://www.thebookseller.com/news/childrens-audiobooks-double-popularity-1086296

Faber, J., & Fonseca, L. M. (2014). How sample size influences research outcomes. *Dental Press Journal of Orthodontics, 19*, 27-29. https://doi.org/10.1590/2176-9451.19.4.027-029.ebo

Farah, M. J. (1994). Neuropsychological inference with an interactive brain: A critique of the "locality" assumption. *Behavioral and Brain Sciences, 17*(1), 43-61. https://doi.org/10.1017/S0140525X00033306

Feil, E. G., Baggett, K., Davis, B., Landry, S., Sheeber, L., Leve, C., & Johnson, U. (2020). Randomized control trial of an internet-based parenting intervention for mothers of infants. *Early Childhood Research Quarterly, 50*, 36-44. https://doi.org/10.1016/j.ecresq.2018.11.003

Feinstein, L. (2003). Inequality in the early cognitive development of British children in the 1970 cohort. *Economica, 70*(277), 73-97. https://doi.org/10.1111/1468-0335.t01-1-00272

Feldman, A. (2007). Validity and quality in action research. *Educational Action Research, 15*(1), 21–32. https://doi.org/10.1080/09650790601150766

Feldman, A., Altrichter, H. Posch, P., & Somekh, B. (2018). *Teachers investigate their work: an introduction to action research across the professions.* (3rd ed.). Routledge.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*(5), 1-185. https://doi.org/10.2307/1166093

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories* (2nd ed.). Brookes Publishing Company.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *CDI-III.* Brookes Publishing Company.

Fergusson, D. M., & Woodward, L. J. (1999). Maternal age and educational and psychosocial outcomes in early adulthood. *The Journal of Child Psychology and Psychiatry and Allied Disciplines, 40*(3), 479-489. https://doi.org/10.1111/1469-7610.00464

Fernald, A. (1985). Four-month-old infants prefer to listen to motherese. *Infant Behavior and Development, 8*(2), 181-195. https://doi.org/10.1016/S0163-6383(85)80005-9

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development, 83*(1), 203-222. https://doi.org/10.1111/j.1467-8624.2011.01692.x

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

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology, 42*(1), 98-116. https://doi.org/10.1037/0012-1649.42.1.98

Field, F. (2010). *The Foundation Years: preventing poor children becoming poor adults*. HM Government. http://www.complexneeds.org.uk/modules/Module-1.1-Understanding-the-child-development-and-difficulties/All/downloads/m01p040d/the_foundation_years_preventing_poor_children_becoming_poor_adults.pdf

Finnegan, J., & Warren, H. (2015). *Ready to read: closing the gap in early language skills so that every child in England can read well.* Save the Children. https://ffteducationdatalab.org.uk/wp-content/uploads/2016/02/Ready_to_Read_England.pdf

Floccia, C., Keren-Portnoy, T., DePaolis, R., Duffy, H., Delle Luche, C., Durrant, S., White, L., Goslin, J., & Vihman, M. (2016). British English infants segment words only with exaggerated infant-directed speech stimuli. *Cognition, 148*, 1-9. https://doi.org/10.1016/j.cognition.2015.12.004

Floor, P., & Akhtar, N. (2006). Can 18-month-old infants learn words by listening in on conversations? *Infancy, 9*(3), 327-339. https://doi.org/10.1207/s15327078in0903_4

Flouri, E. & Malmberg, L. (2012). Father involvement, family poverty and adversity, and young children's behaviour in stable resident two parent families. *Longitudinal and Life Course Studies, 3*(2), 254-267. https://doi.org/10.14301/llcs.v3i2.170

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2021). *Variability and consistency in early language learning: The Wordbank project*. MIT Press.

Fry, A. F., & Hale, S. (1996). Processing Speed, Working Memory, and Fluid Intelligence: Evidence for a Developmental Cascade. *Psychological Science, 7*(4), 237-241. https://doi.org/10.1111/j.1467-9280.1996.tb00366.x

Gajos, J. M., & Beaver, K. M. (2017). Maternal depression and risk for antisocial behaviour in children. *Child & Family Social Work, 22*(1), 349-363. https://doi.org/10.1111/cfs.12247

Gascoigne, M., & Gross, J. (2017). *Talking about a Generation.* The Communication Trust. https://static1.squarespace.com/static/60fe8d94d62921227d256be9/t/61239ecafa9cf435f15611df/1629724365691/tct_talkingaboutageneration_report_online.pdf

Gathercole, S. E., Frankish, C. R., Pickering, S. J., & Peaker, S. (1999). Phonotactic influences on short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*(1), 84-95. https://doi.org/10.1037/0278-7393.25.1.84

Gathercole, S. E., Hitch, G. J., & Martin, A. J. (1997). Phonological short-term memory and new word learning in children. *Developmental Psychology, 33*(6), 966-979. https://doi.org/10.1037/0012-1649.33.6.966

Gentner, D. (2006). Why verbs are hard to learn. In K. Hirsh-Pasek, & R. Golinkoff (Eds.), *Action Meets Word: How Children Learn Verbs* (pp.544-564). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195170009.001.0001

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. In M. Bowerman, & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp.215-256). Cambridge University Press.

Gerhand, S., & Barry, C. (1998). Word frequency effects in oral reading are not merely age-of-acquisition effects in disguise. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(2), 267-283. https://doi.org/10.1037/0278-7393.24.2.267

Gershkoff-Stowe, L., & Hahn, E. R. (2007). Fast mapping skills in the developing lexicon. *Journal of Speech, Language and Hearing Research, 50*(3), 682-697. https://doi.org/10.1044/1092-4388(2007/048)

Gershkoff-Stowe, L., & Hahn, E. R. (2013). Word comprehension and production asymmetries in children and adults. Journal of Experimental Child Psychology, 114, 489–509. https://doi.org/10.1016/j.jecp.2012.11.005

Gibson, T. A., Peña, E. D., & Bedore, L. M. (2012). The relation between language experience and receptive-expressive semantic gaps in bilingual children. *International Journal of Bilingual Education and Bilingualism, 17*(1), 90–110. https://doi.org/10.1080/13670050.2012.743960

Gigerenzer, G. (2015). On the Supposed Evidence for Libertarian Paternalism. *Review of Philosophy and Psychology, 6,* 361–383. https://doi.org/10.1007/s13164-015-0248-1

Gilbert, L., Teravainen, A., Clark, C., & Shaw, S. (2018). *Literacy and life expectancy: An evidence review exploring the link between literacy and life expectancy in England through health and socioeconomic factors*. National Literacy Trust. https://nlt.cdn.ngo/media/documents/National_Literacy_Trust_-_Literacy_and_life_expectancy_report.pdf

Giletta, M., Choukas-Bradley, S., Maes, M., Linthicum, K. P., Card, N. A., & Prinstein, M. J. (2021). A meta-analysis of longitudinal peer influence effects in childhood and adolescence. *Psychological Bulletin, 147*(7), 719-747. https://doi.org/10.1037/bul0000329

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

Girard, L., Pingault, J., Doyle, O., Falissard, B., & Tremblay, R. E. (2016). Developmental associations between conduct problems and expressive language in early childhood: A population-based study. *Journal of Abnormal Child Psychology, 44*, 1033-1043. https://doi.org/10.1007/s10802-015-0094-8

Girolametto, L., Wiigs, M., Smyth, R., Weitzman, E., & Pearce, P. S. (2001). Children with a history of expressive vocabulary delay. *American Journal of Speech-Language Pathology, 10*(4), 358-369. https://doi.org/10.1044/1058-0360(2001/030)

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1*(1), 3-55. https://doi.org/10.1207/s15327817la0101_2

Goldin-Meadow, S., Levine, S. C., Hedges, L. V., Huttenlocher, J., Raudenbush, S. W., & Small, S. L. (2014). New evidence about language and cognitive development based on a longitudinal study: hypotheses for intervention. *American Psychologist, 69*(6), 588-599. https://doi.org/10.1037/a0036886

Goldkuhl, G. (2012). Pragmatism vs interpretivism in qualitative information systems research. *European Journal of Information Systems, 21*, 135-146. https://doi.org/10.1057/ejis.2011.54

Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby) talk to me: The social context of infant-directed speech and its effects on early language acquisition. *Current Directions in Psychological Science, 24*(5), 339-344. https://doi.org/10.1177/0963721415595345

Golinkoff, R. M., Hirsh-Pasek, K., Bloom, L., Smith, L. B., Woodward, A. L., Akhtar, N., Tomasello, M., & Hollich, G. (2000). *Becoming a word learner: A debate on lexical acquisition*. Oxford University Press.

Golinkoff, R. M., & Hirsh-Pasek, K. (2007). Language Development: The View from the Radical Middle. *Proceedings of the 31st Annual Boston University Conference on Language Development, USA*, 1-25.

Golinkoff, R. M., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in Cognitive Sciences, 12*(10), 397-403. https://doi.org/10.1016/j.tics.2008.07.003

Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2019). Language matters: Denying the existence of the 30-million-word gap has serious consequences. *Child Development, 90*(3), 985-992. https://doi.org/10.1111/cdev.13128

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: the case for a developmental lexical principles framework. *Journal of Child Language, 21*(1), 125-155. https://doi:10.1017/S0305000900008692

Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language, 35*(3), 515-531. https://doi:10.1017/S0305000907008641

Goodwyn, S. W., Acredolo, L. P., & Brown, C. A. (2000). Impact of symbolic gesturing on early language development. *Journal of Nonverbal Behavior, 24*, 81-103. https://doi.org/10.1023/A:1006653828895

Great Ormond Street Hospital for Children (n.d.). https://www.gosh.nhs.uk/conditions-and-treatments/procedures-and-treatments/speech-and-language-development-birth-12-months/

Greenwood, C. R., Schnitz, A. G., Carta, J. J., Wallisch, A., & Irvin, D. W. (2020). A systematic review of language intervention research with low-income families: A word gap prevention perspective. *Early Childhood Research Quarterly, 50*, 230-245. https://doi.org/10.1016/j.ecresq.2019.04.001

Greenwood, D. J., & Levin, M. (2006). *Introduction to action research: Social research for social change*. SAGE publications.

Gross, C., Gottsburgsen, A. & Phoenix, A. (2016). Education systems and intersectionality. In: Hadjar, A. & Gross, C. (Eds.), *Education Systems and Inequalities: International Comparisons* (pp.51-71). Policy Press.

Guo, (2023). Cascades in language acquisition: Re-thinking the linear model of development. In C. S. Tamis-Lemonda, & J. J. Lockman (Eds.), *Advances in Child Development and Behavior* (pp.69-107). Academic Press.

Hadley, P. A., Rispoli, M., & Hsu, N. (2016). Toddlers' verb lexicon diversity and grammatical outcomes. *Language, Speech, and Hearing Services in Schools, 47*(1), 44-58. https://doi.org/10.1044/2015_LSHSS-15-0018

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language, 27*(3), 689-705. https://doi:10.1017/S0305000900004414

Hammer, C.S., Morgan, P. Farkas, G., Hillemeier, M., Bitetti, D. & Maczuga, S. (2017). Late Talkers: A Population-Based Study of Risk Factors and School Readiness Consequences. *Journal of Speech, Language, and Hearing Research, 60*(3), 607-626. https://doi.org/10.1044/2016_JSLHR-L-15-0417

Hammersley, M. (1992). *What's wrong with ethnography?* Routledge.

Hansen, P. (2017). What makes a word easy to acquire? The effects of word class, frequency, imageability and phonological neighbourhood density on lexical development. *First Language, 37*(2), 205-225. https://doi.org/10.1177/0142723716679956

Harley, J., Bolton, L., Barton, G., Quincey, A., Gross, J., Wooldridge, J., Reedy, D., Cremin, T., Clements, J., Cain, K., Oakhill, J. V., Coelho, J., Thompson, D., & Gupta, V. (2018). *Why closing the word gap matters.* Oxford language report. https://www.oup.com.cn/test/word-gap.pdf

Harley, T. A. (2001). *The psychology of language: From data to theory* (2nd ed). Psychology Press.

Harley, T. A. (2013). *The psychology of language: From data to theory* (4th ed.). Psychology Press.

Harper, L. J. (2011). Nursery Rhyme Knowledge and Phonological Awareness in Preschool Children. *Journal of Language and Literacy Education, 7*(1), 65-78.

Hart, B., & Risley, T. R. (1995). *Meaningful Differences in the Everyday Experience of Young American Children.* Brookes Publishing Company.

Hartman, K. M., Ratner, N. B., & Newman, R. S. (2017). Infant-directed speech (IDS) vowel clarity and child language outcomes. *Journal of Child Language, 44*(5), 1140-1162. https://doi:10.1017/S0305000916000520

Havron, N., Ramus, F., Heude, B., Forhan, A., Cristia, A., Peyre, H., & EDEN Mother-Child Cohort Study Group. (2019). The effect of older siblings on language development as a function of age difference and sex. *Psychological Science, 30*(9), 1333-1343. https://doi.org/10.1177/0956797619861436

Hayiou-Thomas, M. E., Carroll, J. M., Leavett, R., Hulme, C., & Snowling, M. J. (2017). When does speech sound disorder matter for literacy? The role of disordered speech errors, co-occurring language impairment and

family risk of dyslexia. *Journal of Child Psychology and Psychiatry, 58*(2), 197-205. https://doi.org/10.1111/jcpp.12648

He, A. X., Kon, M., & Arunachalam, S. (2020). Linguistic context in verb learning: Less is sometimes more. *Language Learning and Development, 16*(1), 22-42. https://doi.org/10.1080/15475441.2019.1676751

Heidlage, J. K., Cunningham, J. E., Kaiser, A. P., Trivette, C. M., Barton, E. E., Frey, J. R., & Roberts, M. Y. (2020). The effects of parent-implemented language interventions on child linguistic outcomes: A meta-analysis. *Early Childhood Research Quarterly, 50*, 6-23. https://doi.org/10.1016/j.ecresq.2018.12.006

Heikkinen, H. L. T., Huttunen, R., & Syrjälä, L. (2007). Action research as narrative: five principles for validation. *Educational Action Research, 15*(1), 5–19. https://doi.org/10.1080/09650790601150709

Heymann, P., Heflin, B. H., Baralt, M., & Bagner, D. M. (2020). Infant-directed language following a brief behavioral parenting intervention: The importance of language quality. *Infant Behavior and Development, 58*, 1-9. https://doi.org/10.1016/j.infbeh.2019.101419

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111-127. https://doi.org/10.1207/s15326985ep4102_4

Hilbrink, E. E., Gattis, M. & Levinson, S. C. (2015). Early developmental changes in the timing of turn-taking: a longitudinal study of mother–infant interaction. *Frontiers in Psychology, 6,* 246-257. https://doi.org/10.3389/fpsyg.2015.01492

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009). Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition? *Psychological Science, 20*(6), 729-739. https://doi.org/10.1111/j.1467-9280.2009.02365.x

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., Yust, P. K., & Suma, K. (2015). The contribution of early communication quality to low-income children's language success. *Psychological Science, 26*(7), 1071-1083. https://doi.org/10.1177/0956797615581493

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*(5), 1368-1378. https://doi.org/10.1111/1467-8624.00612

Hoff, E. (2011). *Research Method in Child Language: A Practical Guide.* Wiley Blackwell.

Hoff, E., & Tian, C. (2005). Socioeconomic status and cultural influences on language. *Journal of Communication Disorders, 38*(4), 271-278. https://doi.org/10.1016/j.jcomdis.2005.02.003

Hoff-Ginsberg, E. (1986). Function and structure in maternal speech: Their relation to the child's development of syntax. *Developmental Psychology, 22*(2), 155-163. https://doi.org/10.1037/0012-1649.22.2.155

Hoff-Ginsberg, E. (1998). The relation of birth order and socioeconomic status to children's language experience and language development. *Applied Psycholinguistics, 19*(4), 603-629. https://doi.org/10.1017/S0142716400010389

Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., Hennon, E., Rocroi, C., & Bloom, L. (2000). Breaking the language barrier: An emergentist coalition model for the origins of word learning. *Monographs of the Society for Research in Child Development, 65*(3), i-135.

Horst, J. S., Parsons, K. L., & Bryan, N. M. (2011). Get the story straight: Contextual repetition promotes word learning from storybooks. *Frontiers in Psychology, 2*, 1-11. https://doi.org/10.3389/fpsyg.2011.00017

Horvath, S., Kueser, J. B., Kelly, J., & Borovsky, A. (2022). Difference or delay? Syntax, semantics, and verb vocabulary development in typically developing and late-talking toddlers. *Language Learning and Development, 18*(3), 352-376. https://doi.org/10.1080/15475441.2021.1977645

Hoyle, S.M. & Adger, C.T. (Eds.). (1998). *Kids Talk: Strategic Language Use in Later Childhood.* Oxford University Press.

Huang, Y. T., Leech, K., & Rowe, M. L. (2017). Exploring socioeconomic differences in syntactic development through the lens of real-time processing. *Cognition, 159*, 61-75. https://doi.org/10.1016/j.cognition.2016.11.004

Hudry, K., Leadbitter, K., Temple, K., Slonims, V., McConachie, H., Aldred, C., et al. (2010). Preschoolers with autism show greater impairment in receptive compared with expressive language abilities. *International Journal of Language and Communication Disorders, 45*(6), 681–690. https://doi.org/10.3109/13682820903461493

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: relation to language input and gender. *Developmental Psychology, 27*(2), 236-248. https://doi.org/10.1037/0012-1649.27.2.236

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology, 45*(3), 337-374. https://doi.org/10.1016/S0010-0285(02)00500-5

Huttenlocher, J., Vasilyeva, M., Waterfall, H. R., Vevea, J. L., & Hedges, L. V. (2007). The varieties of speech to young children. *Developmental Psychology, 43*(5), 1062-1083. https://doi.org/10.1037/0012-1649.43.5.1062

Huttenlocher, J., Waterfall, H., Vasilyeva, M., Vevea, J., & Hedges, L. V. (2010). Sources of variability in children's language growth. *Cognitive Psychology, 61*(4), 343-365. https://doi.org/10.1016/j.cogpsych.2010.08.002

I CAN Charity (n.d.). https://speechandlanguage.org.uk/help-for-families/ages-and-stages/*Infant Feeding Survey - UK, 2010.* (2012, https://digital.nhs.uk/data-and-information/publications/statistical/infant-feeding-survey/infant-feeding-survey-uk-2010

Iverson, J. & Goldin-Meadow, S. (2005). Gesture paves the way for language development. *Psychological Science, 16*, 367-371. https://doi.org/10.1111/j.0956-7976.2005.01542.x

Jenson, S. S. (2023). The timing of parental unemployment and children's academic achievement. *Advances in Life Course Research, 57,* 1-13. https://doi.org/10.1016/j.alcr.2023.100557

Jeong, J., McCoy, D. C., & Fink, G. (2017). Pathways between paternal and maternal education, caregivers' support for learning, and early child development in 44 low-and middle-income countries. *Early Childhood Research Quarterly, 41*, 136-148. https://doi.org/10.1016/j.ecresq.2017.07.001

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language, 44*(4), 548-567. https://doi.org/10.1006/jmla.2000.2755

Jones, G., Cabiddu, F., Andrews, M., & Rowland, C. (2021). Chunks of phonological knowledge play a significant role in children's word learning and explain effects of neighborhood size, phonotactic probability, word frequency and word length. *Journal of Memory and Language, 119*, 1-13. https://doi.org/10.1016/j.jml.2021.104232

Jones, G., Cabiddu, F., Barrett, D.J.K., Castro, A. & Lee, B. (2023). How the characteristics of words in child-directed speech differ from adult-directed speech to influence children's productive vocabularies. *First Language, 43*(3), 253-282. https://doi.org/10.1177/01427237221150070

Jones, G., & Rowland, C. F. (2017). Diversity not quantity in caregiver speech: Using computational modeling to isolate the effects of the quantity and the diversity of the input on vocabulary growth. *Cognitive Psychology, 98*, 1-21. https://doi.org/10.1016/j.cogpsych.2017.07.002

Jones, S. D., & Brandt, S. (2019a). Do children really acquire dense neighbourhoods? *Journal of Child Language, 46*(6), 1260-1273. https://doi.org/10.1017/S0305000919000473

Jusczyk, P. W. (2000). *The Discovery of Spoken Language.* The MIT Press.

Jusczyk, P. W., Houston, D. M., & Newsome, M. (1999). The beginnings of word segmentation in English-learning infants. *Cognitive Psychology, 39*(3-4), 159-207. https://doi.org/10.1006/cogp.1999.0716

Just, J. (2018). *Assessing the predictive value of the UK-CDI for early identification of developmental language delay* [Doctoral dissertation, University of Lincoln]. University of Lincoln Repository. https://repository.lincoln.ac.uk/articles/thesis/Assessing_the_predictive_value_of_the_UK-CDI_for_early_identification_of_developmental_language_delay/24325987

Kachergis, G., Hembacher, E., Cristiano, V., Zhang, V., & Frank, M. C. (2020). *Understanding the impacts of video-guided activities on parent-child interaction.* PsyArXiv. https://doi.org/10.31234/osf.io/dmeza

Kaushik, V., & Walsh, C. A. (2019). Pragmatism as a research paradigm and its implications for social work research. *Social Sciences, 8*(9), 1-17. https://doi.org/10.3390/socsci8090255

Kelly, M. H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review, 99*(2), 349–364. https://doi.org/10.1037/0033-295X.99.2.349

Kennedy, G. (2014). *An introduction to corpus linguistics*. Routledge. https://doi.org/10.4324/9781315843674

Kilgour, C., & Fleming, V. (2000). An action research inquiry into a health visitor parenting programme for parents of pre-school children with behaviour problems. *Journal of Advanced Nursing, 32*(3), 682-688. https://doi.org/10.1046/j.1365-2648.2000.01528.x

Kim, J., Gilbert, J., Yu, Q., & Gale, C. (2021). Measures Matter: A Meta-Analysis of the Effects of Educational Apps on Preschool to Grade 3 Children's Literacy and Math Skills. *AERA Open*, *7*(1), 1-19. https://doi.org/10.1177/23328584211004183

Kimbrough Oller, D. (2010). All-day recordings to investigate vocabulary development: A case study of a trilingual toddler. *Communication Disorders Quarterly, 31*(4), 213-222. https://doi.org/10.1177/1525740109358628

Kolbe, D. A. (1984). *Experiential learning: experience as the source of learning and development.* Prentice Hall.

Kowalsky, M. (2008). *Right here, Right Now: Children's use of Adverbs* (Publication No. 3331498) [Doctoral dissertation, University of Minnesota]. ProQuest Dissertations & Theses Global.

Kroese, F. M., Marchiori, D. R., & De Ridder, D. T. (2016). Nudging healthy food choices: a field experiment at the train station. *Journal of Public Health, 38*(2), e133-e137. https://doi.org/10.1093/pubmed/fdv096

Kuhn, T. S. (1962). *The structure of scientific revolutions.* University of Chicago Press.

Kwiatkowski, B. (2019). The symbiotic roles of action research, lesson study and learning study seen in a social–emotional intervention for males with behavioural needs. *Educational Action Research, 27*(4), 613–630. https://doi-org.ntu.idm.oclc.org/10.1080/09650792.2019.1570858

Laing, C. E., & Bergelson, E. (2024). Analyzing the effect of sibling number on input and output in the first 18 months. *Infancy, 29*(2), 175-195. https://doi.org/10.1111/infa.12578

Lamb, M. E. (Ed.). (2004). *The role of the father in child development*. (4th ed.). John Wiley & Sons, Inc.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104*(2), 211-240. https://doi.org/10.1037/0033-295X.104.2.211

La Prairie, D. (2010). Effect Of Prolonged Pacifier Use On Speech Articulation. [Master's thesis: Eastern Illinois University]. Eastern Illinois University Institutional Repository. http://thekeep.eiu.edu/theses/100

Larson, A. L., An, Z. G., Wood, C., Uchikoshi, Y., Cycyk, L. M., Scheffner Hammer, C., Escobar, K., & Roberts, K. (2020). Social validity in early language interventions for dual language learners: A systematic review of the literature. *Topics in Early Childhood Special Education, 40*(1), 39-51. https://doi.org/10.1177/0271121419901289

Lau, Y. L., Loi, C. K., & bin Abdullah, M. N. A. (2021). The Historical Development Of The Study Of Broca's Aphasia. *Malang Neurology Journal, 7*(2), 125-128. : httpps://doi.org/10.21776/ub.mnj.2020.007.02.8

Law, J., Boyle, J., Harris, F., & Harkness, A. (1998). Child health surveillance: screening for speech and language delay. ePrints. https://eprints.ncl.ac.uk/156218.

Law, J., Charlton, J., & Asmussen, K. (2017b). *Language as a Child Wellbeing Indicator.* Early Intervention Foundation. https://www.eif.org.uk/report/language-as-a-child-wellbeing-indicator

Law, J., Charlton, J., Dockrell, J., Gascoigne, M., McKean, C., & Theakston, A. (2017a). *Early language development: Needs, provision and intervention for pre-school children from socio-economically disadvantaged backgrounds*. Education Endowment Foundation. https://www.yhphnetwork.co.uk/media/1494/early-language-development-needs-provision-and-intervention-for-preschool-children-from-socio-economically-disadvantaged-backgrounds-phe-2017.pdf

Law, J., Clegg, J., Rush, R., Roulstone, S., & Peters, T. J. (2019). Association of proximal elements of social disadvantage with children's language development at 2 years: an analysis of data from the Children in Focus (CiF) sample from the ALSPAC birth cohort. *International Journal of Language & Communication Disorders, 54*(3), 362-376. https://doi.org/10.1111/1460-6984.12442

Lee, D. (2018, July 10) Audiobooks vs Reading: The Rules Are, There Are No Rules. *Book Riot.* https://bookriot.com/audiobooks-vs-reading/

Lee, J. A. C. & Al Otaiba, S. (2015). Socioeconomic and gender group differences in early literacy skills: a multiple-group confirmatory factor analysis approach. *Educational Research and Evaluation, 21*(1), 40–59. https://doi.org/10.1080/13803611.2015.1010545

Leigh, P., Nievar, M. A., & Nathans, L. (2011). Maternal sensitivity and language in early childhood: A test of the transactional model. *Perceptual and Motor Skills, 113*(1), 281-299. https://doi.org/10.2466/10.17.21.28.PMS.113.4.281-299

Lewis, M. L., & Frank, M. C. (2016). The length of words reflects their conceptual complexity. *Cognition, 153*, 182-195. https://doi.org/10.1016/j.cognition.2016.04.003

Lindsay, L., Gambi, C. & Rabagliati, H. (2019). Preschoolers Optimize the Timing of Their Conversational Turns Through Flexible Coordination of Language Comprehension and Production. *Psychological Science, 30*(4), 504-515. https://doi.org/10.1177/0956797618822802

Liszkowski, U., & Tomasello, M. (2011). Individual differences in social, cognitive, and morphological aspects of infant pointing. *Cognitive Development, 26*(1), 16-29. https://doi.org/10.1016/j.cogdev.2010.10.001

Locke, A., Ginsborg, J., & Peers, I. (2002). Development and disadvantage: Implications for the early years and beyond. *International Journal of Language & Communication Disorders, 37*(1), 3-15. https://doi.org/10.1080/13682820110089911

Loizou, E. (2013). Empowering parents through an action research parenting program. *Action Research, 11*(1), 73-91. https://doi.org/10.1177/147675031347631

Lonigan, C. J., & Whitehurst, G. J. (1998). Relative efficacy of parent and teacher involvement in a shared-reading intervention for preschool children from low-income backgrounds. *Early Childhood Research Quarterly, 13*(2), 263-290. https://doi.org/10.1016/S0885-2006(99)80038-6

Lust, B.C. & Blume, M. (2016). *Research methods in language acquisition: Principles, procedures, and practices.* De Gruyter, Inc. https://doi.org/10.1515/9783110415339

MacRoy-Higgins, M., Shafer, V. L., Fahey, K. J., & Kaden, E. R. (2016). Vocabulary of toddlers who are late talkers. *Journal of Early Intervention, 38*(2), 118-129. https://doi.org/10.1177/1053815116637620

MacWhinney, B. (2000). The CHILDES Project: Tools for Analyzing Talk. (3rd ed.). Lawrence Erlbaum Associates Inc. https://doi.org/10.21415/T5G10R

MacWhinney, B. (2000). The CHILDES Project: Tools for Analyzing Talk. (3rd ed.). Lawrence Erlbaum Associates Inc. https://doi.org/10.21415/3mhn-0z89

Maekawa, J., & Storkel, H. L. (2006) Individual differences in the influence of phonological characteristics on expressive vocabulary development by young children. *Journal of Child Language, 33*(3), 439-459. https://doi.org/10.1017/S0305000906007458

Maier, M., Bartoš, F., Stanley, T. D., Shanks, D. R., Harris, A. J., & Wagenmakers, E. (2022). No evidence for nudging after adjusting for publication bias. *Proceedings of the National Academy of Sciences, 119*(31), 1-2. https://doi.org/10.1073/pnas.2200300119

Malaspina, D., Reichenberg, A., Weiser, M., Fennig, S., Davidson, M., Harlap, S., Wolitzky, R., Rabinowitz, J., Susser, E., & Knobler, H. Y. (2005). Paternal age and intelligence: implications for age-related genomic changes in male germ cells. *Psychiatric Genetics, 15*(2), 117-125.

Malin, J. L., Karberg, E., Cabrera, N. J., Rowe, M., Cristaforo, T., & Tamis-LeMonda, C. S. (2012). Father–toddler communication in low-income families: The role of paternal education and depressive symptoms. *Family Science, 3*(3-4), 155-163. https://doi.org/10.1080/19424620.2012.779423

Mampe, B., Friederici, A. D., Christophe, A., & Wermke, K. (2009). Newborns' cry melody is shaped by their native language. *Current Biology, 19*(23), 1994-1997. https://doi.org/10.1016/j.cub.2009.09.064

Manz, P. H., Hughes, C., Barnabas, E., Bracaliello, C., & Ginsburg-Block, M. (2010). A descriptive review and meta-analysis of family-based emergent literacy interventions: To what extent is the research applicable to low-income, ethnic-minority or linguistically-diverse young children? Early Childhood Research Quarterly, 25(4), 409-431. https://doi.org/10.1016/j.ecresq.2010.03.002

Marchman, V. A., Adams, K. A., Loi, E. C., Fernald, A., & Feldman, H. M. (2016). Early language processing efficiency predicts later receptive vocabulary outcomes in children born preterm. *Child Neuropsychology, 22*(6), 649-665. https://doi.org/10.1080/09297049.2015.1038987

Marchman, V. A., Dale, P. S., & Fenson, L. (2023). *The MacArthur-Bates Communicative Development Inventories: User's Guide and Technical Manual.* (3rd ed.). Brookes Publishing Co.

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science 283*(5398)*,* 77–80. https://doi.org/10.1126/science.283.5398.77

Mare, R. D., & Tzeng, M. (1989). Fathers' ages and the social stratification of sons. *American Journal of Sociology, 95*(1), 108-131. https://doi.org/10.1086/229215

Markman, E. M. (1989). *Categorization and Naming in Children: Problems of Induction*. MIT Press.

Mason-Apps, E., Stojanovik, V., & Houston-Price, C. (2011, August 17-21). Early Word Segmentation in Typically Developing Infants and Infants with Down Syndrome: A Preliminary Study [Conference presentation]. International Conference of Phonetic Sciences, Hong Kong.

Matthews, D., Lieven, E., Theakston, A., & Tomasello, M. (2009). Pronoun co-referencing errors: Challenges for generativist and usage-based accounts. *Cognitive Linguistics, 20,* 599–626. https://doi.org/10.1515/COGL.2009.026

Matthews, J. S., Ponitz, C. C., & Morrison, F. J. (2009). Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology, 101*(3), 689-674. https://doi.org/10.1037/a0014240

McAteer, M. (2013). *Action Research in Education.* Sage Publications Ltd. https://doi.org/10.4135/9781473913967

McArdle, K. L. (2004). In-powering spaces: a cooperative inquiry with young women in management [Unpublished doctoral dissertation]. University of Bath.

McCarthy, P. M., & Jarvis, S. (2007). VOCD: A theoretical and empirical evaluation. *Language Testing, 24*(4), 459-572. https://doi.org/10.1177/0265532207080767

McEnery, T., & Gabrielatos, C. (2006). English corpus linguistics. In B. Aarts, & A. McMahon (Eds.). *The Handbook of English Linguistics* (pp. 33-71). Blackwell Publishing Ltd. https://doi.org/10.1002/9780470753002

McGillion, M., Herbert, J. S., Pine, J., Vihman, M., DePaolis, R., Keren-Portnoy, T., & Matthews, D. (2017). What paves the way to conventional language? The predictive value of babble, pointing, and socioeconomic status. *Child Development, 88*(1), 156-166. https://doi.org/10.1111/cdev.12671

McKernan, J. (1991). Action inquiry: Studied enactment. In E. C. Short (Ed.). *Forms of Curriculum Inquiry* (pp.309-326). New York Press.

McNiff, J. (2002). *Action research for professional development*. Hyde Publications.

McNiff, J., & Whitehead, J. (2002). *Action research in organisations.* Routledge.

McNiff, J., & Whitehead, J. (2005). *Action research for teachers: A practical guide*. Routledge. https://doi.org/10.4324/9780203462393

Meints, K., & Fletcher, K. (2011). Lincoln University Babylab Toddler Communicative Development Inventory, University of Lincoln.

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*(2), 270-291. https://doi.org/10.1037/a0028228

Melhuish, E. C. (2016). Provision of quality early childcare services: synthesis report. Birbeck Institutional Research Online. http://dx.doi.org/10.2767/757173

Melhuish, E., Sylva, K., Sammons, P., Siraj-Blatchford, I., & Taggart, B. (2001). Home Learning Environment Index (1st ed.). Edward Melhuish OBE.

Melnikoff, D. E., & Bargh, J. A. (2018). The mythical number two. *Trends in Cognitive Sciences, 22*(4), 280-293. https://doi.org/10.1016/j.tics.2018.02.001

Merriman, W. E., Bowman, L. L., & MacWhinney, B. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development, 54*(3/4), i-129. https://doi.org/10.2307/1166130

Mertens, S., Herberz, M., Hahnel, U. J., & Brosch, T. (2022). The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences, 119*(1), 1-10. https://doi.org/10.1073/pnas.2107346118

Messenger, K., Yuan, S., & Fisher, C. (2015). Learning verb syntax via listening: New evidence from 22-month-olds. *Language Learning and Development, 11*(4), 356-368. https://doi.org/10.1080/15475441.2014.978331

Miller, L. (2010, June 30) “War and Peace" made easy. *Salon.* https://www.salon.com/2010/06/30/audiobooks_2/

Milner IV, H. R. (2013). Rethinking achievement gap talk in urban education. *Urban Education, 48*(1), 3-8. https://doi.org/10.1177/0042085912470417

Moeller, M. P., Hoover, B., Putman, C., Arbataitis, K., Bohnenkamp, G., Peterson, B., Wood, S., Lewis, D., Pittman, A., & Stelmachowicz, P. (2007). Vocalizations of infants with hearing loss compared with infants with normal hearing: Part I–phonetic development. *Ear and Hearing, 28*(5), 605-627. https://doit.org/10.1097/AUD.0b013e31812564ab

Mol, S. E., Bus, A. G., de Jong, M. T., & Smeets, D. J. H. (2008). Added Value of Dialogic Parent–Child Book Readings: A Meta-Analysis. *Early Education and Development, 19*(1), 7–26. https://doi.org/10.1080/10409280701838603

Mooney, J., Winter, K., & Connolly, P. (2016). Effects of a book gifting programme on literacy outcomes for foster children: A randomised controlled trial evaluation of the Letterbox Club in Northern Ireland. *Children and Youth Services Review, 65,* 1-8. https://doi.org/10.1016/j.childyouth.2016.03.009

Moore, M. R., & Brooks-Gunn, J. (2002). Adolescent parenthood. In M. H. Bornstein (Ed.). *Handbook of Parenting: being and becoming a parent* (pp.173-214). Lawrence Erlbaum Associates Inc.

Morgan, D. L. (2014). Pragmatism as a paradigm for social research. *Qualitative Inquiry, 20*(8), 1045-1053. https://doi.org/10.1177/1077800413513733

Morgan, L., & Wren, Y. E. (2018). A systematic review of the literature on early vocalizations and babbling patterns in young children. *Communication Disorders Quarterly, 40*(1), 3-14. https://doi.org/10.1177/1525740118760215

Moss, G., & Washbrook, L. (2016). *Understanding the gender gap in literacy and language development*. University of Bristol Graduate School. https://discovery.ucl.ac.uk/id/eprint/10068374/1/Understanding%20the%20Gender%20Gap%20working%20paper.pdf

Moutsiana, C., Garrett, N., Clarke, R. C., Lotto, R. B., Blakemore, S., & Sharot, T. (2013). Human development of the ability to learn from bad news. *Proceedings of the National Academy of Sciences, 110*(41), 16396-16401. https://doi.org/10.1073/pnas.1305631110

Muhinyi, A., Hesketh, A., Stewart, A. J., & Rowland, C. F. (2020). Story choice matters for caregiver extra-textual talk during shared reading with preschoolers. *Journal of Child Language, 47*(3), 633-654. https://doi.org/10.1017/S0305000919000783

Department for Education (2018, April 30) *Multi-million fund to boost children's early language skills.* [Press release]. https://www.gov.uk/government/news/multi-million-fund-to-boost-childrens-early-language-skills#:~:text=A%20new%20%C2%A35million%20scheme,reading%20and%20singing%20nursery%20rhymes

Muñez, D., Bull, R., & Lee, K. (2022). Maternal education and siblings: Agents of cognitive development in kindergarten. *Developmental Science, 25*(4), 1-15. https://doi.org/10.1111/desc.13218

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17*(2), 357-374. https://doi.org/10.1017/S0305000900013817

Naigles, L. R., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of Child Language, 25*(1), 95-120. https://doi.org/10.1017/S0305000997003358

Neville, H. J., Stevens, C., Pakulak, E., Bell, T. A., Fanning, J., Klein, S., & Isbell, E. (2013). Family-based training program improves brain function, cognition, and behavior in lower socioeconomic status preschoolers. *Proceedings of the National Academy of Sciences, 110*(29), 12138-12143. https://doi.org/10.1073/pnas.130443711

Nightingale, R. (2019). *All Party Parliamentary Group on Social Mobility-EARLY YEARS.* Sutton Trust. https://policycommons.net/artifacts/2612023/all-party-parliamentary-group-on-social-mobility/3634578/

Noble, C., Cameron-Faulkner, T., Jessop, A., Coates, A., Sawyer, H., Taylor-Ims, R., & Rowland, C. F. (2020). The Impact of Interactive Shared Book Reading on Children's Language Skills: A Randomized Controlled Trial. *Journal of Speech, Language, and Hearing Research, 63*(6), 1878-1897. https://doi.org/10.1044/2020_JSLHR-19-00288

Noccetti, S., Tribushinina, E., & Voeikova, M. D. (Eds.). (2015). *Semantics and morphology of early adjectives in first language acquisition*. Cambridge Scholars Publishing.

Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E., Vamvakas, G., & Pickles, A. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: Evidence from a population study. *Journal of Child Psychology and Psychiatry, 57*(11), 1247-1257. https://doi.org/10.1111/jcpp.12573

Oancea, A., & Furlong, J. (2007). Expressions of excellence and the assessment of applied and practice-based research. In A. Oancea, & J. Furlong (Eds.), *Assessing quality in applied and practice-based research in education* (pp. 11-30). Routledge.

Ofcom. (2023). *Children and Parents: Media Use and Attitudes.* https://www.ofcom.org.uk/__data/assets/pdf_file/0027/255852/childrens-media-use-and-attitudes-report-2023.pdf

Office for National Statistics. (2017). *Nowcasting household income in the UK: financial year ending 2017.* UK Statistics Authority.

https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/nowcastinghouseholdincomeintheuk/financialyearending2017

Office for National Statistics. (2019). *Families and households in the UK: 2019.* UK Statistics Authority. https://www.ons.gov.uk/releases/familiesandhouseholdsintheuk2019

Office for National Statistics. (2017). *Birth Characteristics in England and Wales: 2017.* UK Statistics Authority. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthcharacteristicsinenglandandwales/2017#average-ages-of-mothers-and-fathers-of-all-babies-have-continued-to-rise

Office of National Statistics. (2023). *UK labour market.* UK Statistics Authority. https://www.ons.gov.uk/releases/uklabourmarketmay2023

Office of National Statistics. (2018). *Nowcasting household income in the UK.* UK Statistics Authority. https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/datasets/nowcastinghouseholdincomeintheuk

Office for National Statistics. (2021). *Birth Characteristics in England and Wales: 2021.* UK Statistics Authority. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthcharacteristicsinenglandandwales/2021

Office for National Statistics. (2022). *Births in England and Wales: 2022 (refreshed populations).* UK Statistics Authority.https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/livebirths/bulletins/birthsummarytablesenglandandwales/2022refreshedpopulations#:~:text=The%20total%20fertility%20rate%20(TFR,where%20the%20fertility%20rate%20increased.

Olson, H., Ozernov-Palchik, O., Arechiga, X. M., & Gabrieli, J. D. E. (2024). *Remote text-supplemented audiobook intervention improves vocabulary knowledge in third and fourth-grade students*. OSF. file:///C:/Users/PSY3PHELAH/Downloads/Preprint_11Mar2024_Olson_Ozernov-Palchik%20(1).pdf

Osman, M. (2022, July 29). Nudge theory doesn't work after all, says new evidence review – but it could still have a future. The Conversation. https://theconversation.com/nudge-theory-doesnt-work-after-all-says-new-evidence-review-but-it-could-still-have-a-future-187635

Ota, M., & Skarabela, B. (2018). Reduplication facilitates early word segmentation. *Journal of Child Language, 45*(1), 204-218. https://doi.org/10.1017/S0305000916000660

Pace, A., Hirsh-Pasek, K., & Golinkoff, R. M. (2016). How high quality language environments create high quality learning environments. In N. K. Lesaux, & S. M. Jones (Eds.), *The Leading Edge of Early Childhood Education* (pp. 45-66). Harvard Education Press.

Page, J., Cock, M. L., Murray, L., Eadie, T., Niklas, F., Scull, J., & Sparling, J. (2019). An Abecedarian Approach with Aboriginal Families and Their Young Children in Australia: Playgroup Participation and Developmental Outcomes. *International Journal of Early Childhood, 51,* 233–250. https://doi.org/10.1007/s13158-019-00246-3

Pan, P. C. D. B. A., Perlmann, R. Y. & Snow, C. E. (1999). Food for thought: Dinner table as a context for observing parent-child discourse. In L. Menn, & N. B. Ratner (Eds.), *Methods for studying language production* (pp. 209-228). Psychology Press.

Pancsofar, N., & Vernon-Feagans, L. (2010). Fathers' early contributions to children's language development in families from low-income rural communities. *Early Childhood Research Quarterly, 25*(4), 450-463. https://doi.org/10.1016/j.ecresq.2010.02.001

Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. *Health Services Research, 34*(5 Pt 2), 1189-1208.

Pérez Granados, L., Alcaraz-Salarirche, N., Fernández Navas, M., & Postigo Fuentes, A. Y. (2022). Action research through lesson study: a space for learning in initial teacher training. *Educational Action Research, 32*(1), 7–24. https://doi.org/10.1080/09650792.2022.2108473

Perez-Pereira, M. (2021). Prevalence of Language Delay among Healthy Preterm Children, Language Outcomes and Predictive Factors. *Children, 8*(4), 1-20. https://doi.org/10.3390/children8040282

Perkins, S. C., Finegood, E. D., & Swain, J. E. (2013). Poverty and language development: Roles of parenting and stress. *Innovations in Clinical Neuroscience, 10*(4), 10-19.

Perry, L. K., Perlman, M., & Lupyan, G. (2015). Iconicity in English and Spanish and its relation to lexical category and age of acquisition. *PloS One, 10*(9), 1-17. https://doi.org/10.1371/journal.pone.0137147

Peter, M. S., Durrant, S., Jessop, A., Bidgood, A., Pine, J. M., & Rowland, C. F. (2019). Does speed of processing or vocabulary size predict later language growth in toddlers? *Cognitive Psychology, 115*, 1-25. https://doi.org/10.1016/j.cogpsych.2019.101238

Pillinger, C., & Wood, C. (2014). Pilot study evaluating the impact of dialogic reading and shared reading at transition to primary school: early literacy skills and parental attitudes. *Literacy, 48*(3), 155-163. https://doi.org/10.1111/lit.12018

Pine, J. M. (1992). Maternal style at the early one-word stage: Re-evaluating the stereotype of the directive mother. *First Language, 12*(35), 169-186. https://doi.org/10.1177/0142723792012035

Pine, J. M. (1995). Variation in vocabulary development as a function of birth order. *Child Development, 66*(1), 272-281. https://doi.org/10.1111/j.1467-8624.1995.tb00870.x

Pleck, J. H., & Masciadrelli, B. P. (1997). Paternal involvement: Levels, sources, and consequences. In M. E. Lamb (Ed.), *The Role of the Father in Child Development* (pp.66-103). John Wiley & Sons, Inc.

Plomin, R., & Dale, P. S. (2014). Genetics and early language development: A UK study of twins. In D. V. M. Bishop, & L. B. Leonard (Eds.), *Speech and language impairments in children* (pp. 35-51). Psychology Press.

Poet, H., Sharp, C., Garry, J., Harland, J., Keightley, G., Kirkup, C., & Martin, K. (2018) *Assessment without levels: qualitative research.* Department for Education. https://assets.publishing.service.gov.uk/media/5f6385f48fa8f510677710cd/NFER_AWL_report.pdf

Polka, L., & Sundara, M. (2003, August 3-9). *Word segmentation in monolingual and bilingual infant learners of English and French* [Conference presentation]. International Conference of Phonetic Sciences, Barcelona.

Pollard, T., Coote, A. Ewart-Biggs, H., Stephens, T., & Sandher, J. (2023). *A Fair Start for All: A universal basic services approach to early education and care.* New Economics Foundation. https://neweconomics.org/uploads/files/Early-years-education-childcare-report-web.pdf

Poulter, S. (2010, April 7). UK: Ten times more stay-at-home dads than 10 years ago. *Aviva.* https://www.aviva.com/newsroom/news-releases/2010/04/uk-ten-times-more-stay-at-home-dads-than-10-years-ago-6344/

Prime, H., Pauker, S., Plamondon, A., Perlman, M., & Jenkins, J. (2014). Sibship size, sibling cognitive sensitivity, and children's receptive vocabulary. *Pediatrics, 133*(2), e394-e401. https://doi.org/10.1542/peds.2012-2874

Puglisi, M. L., Hulme, C., Hamilton, L. G., & Snowling, M. J. (2017). The home literacy environment is a correlate, but perhaps not a cause, of variations in children's language and literacy development. *Scientific Studies of Reading, 21*(6), 498-514. https://doi.org/10.1080/10888438.2017.1346660

Pullum, G. K., & Scholz, B. C. (2002). Empirical assessment of stimulus poverty arguments. *The Linguistic Review, 19*(1-2), 9-50. https://doi.org/10.1515/tlir.19.1-2.9

Purser, H. R. M., Van Herwegen, J., & Thomas, M. S. C. (2020). The development of children's comprehension and appreciation of riddles. *Journal of Experimental Child Psychology, 189*, 1-12. https://doi.org/10.1016/j.jecp.2019.104709

Quine, W. V. O. (1960). *Word and object*. MIT press.

Rabagliati, H., Ferguson, B., & Lew-Williams, C. (2019). The profile of abstract rule learning in infancy: Meta-analytic and experimental evidence. *Developmental Science, 22*(1)*,* 1-18. https://doi.org/10.1111/desc.12704

Ramlal, A., & Augustin, D. S. (2019). Engaging students in reflective writing: an action research project. *Educational Action Research, 28*(3), 518–533. https://doi-org.ntu.idm.oclc.org/10.1080/09650792.2019.1595079

Reason, P., & Bradbury, H. (Eds.). (2001). *Handbook of action research: Participative inquiry and practice*. Sage Publications.

Reece, E., Sparks, A., & Leyva, D. (2010). A Review of parent interventions for preschool children's language and emergent literacy. *Journal of Early Childhood Literacy, 10*(1), 97-117. https://doi.org/10.1177/1468798409356987

Renfrew, C. (1995). *Word Finding Vocabulary Test* (4th ed.). Speechmark Publishing.

Rescorla, L. (2009) Age 17 Language and Reading Outcomes in Late-Talking Toddlers: Support for a Dimensional Perspective on Language Delay. *Journal of Speech, Language, and Hearing Research, 52*(1), 16-30. https://doi.org/10.1044/1092-4388(2008/07-0171)

Rice, M. L., Oetting, J. B., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of Input Effects on Word Comprehension of Children With Specific Language Impairment. *Journal of Speech, Language and Hearing Research, 37*(1), 106-122. https://doi.org/10.1044/jshr.3701.106

Ridge, K. E., Weisberg, D. S., Ilgaz, H., Hirsh-Pasek, K .A., & Golinkoff, R .M. (2015). Supermarket Speak: Increasing Talk Amon Low-Socioeconomic Status Families. *Mind, Brain, and Education, 9*(3), 127-135. https://doi.org/10.1111/mbe.12081

Riding, P., Fowell, S., & Levy, P. (1995). An action research approach to curriculum development. *Information Research, 1*(1).

Roberts, M. Y., & Kaiser, A. P. (2011). The effectiveness of parent-implemented language interventions: A meta-analysis. *American Journal of Speech-Language Pathology, 20*(3), 180-199. https://doi.org/10.1044/1058-0360(2011/10-0055)

Rogowsky, B. A., Calhoun, B. M., & Tallal, P. (2016). Does modality matter? The effects of reading, listening, and dual modality on comprehension. *Sage Open, 6*(3). https://doi.org/10.1177/2158244016669550

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(6), 906-914. https://doi.org/10.1002/wcs.78

Rothbauer, P. (2008). Triangulation. In L. M. Given (Ed.), *The SAGE Encyclopaedia of Qualitative Research Methods* (pp.892-894). Sage Publications.

Roulstone, S., Law, J., Rush, R., Clegg, J., & Peters, T. (2011). *Investigating the role of language in children's early educational outcomes.* Department for Education. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/181549/DFE-RR134.pdf

Roulstone, S., Wren, Y., Bakopoulou, I., Goodlad, S., & Lindsay, G. (2012). *Exploring interventions for children and young people with speech, language and communication needs: A study of practice.* Department for Education. https://www.researchgate.net/profile/Geoff-Lindsay/publication/265906594_Exploring_interventions_for_children_and_young_people_with_speech_language_and_communication_needs_A_study_of_practice/links/5420688d0cf203f155c591dd/Exploring-interventions-for-children-and-young-people-with-speech-language-and-communication-needs-A-study-of-practice.pdf

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development, 83*(5), 1762-1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x

Rowe, M. L. (2018). Understanding socioeconomic differences in parents' speech to children. *Child Development Perspectives, 12*(2), 122-127. https://doi.org/10.1111/cdep.12271

Rowe, M. L., & Goldin-Meadow, S. (2009). Differences in early gesture explain SES disparities in child vocabulary size at school entry. *Science, 323*(5916), 951-953. https://doi.org/10.1126/science.1167025

Rowe, M.L., Turco, R.G. & Blatt, J.H. (2021). Can interactive apps promote parent-child conversations? *Journal of Applied Developmental Psychology, 76,* 1-11. https://doi.org/10.1016/j.appdev.2021.101326

Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences, 112*(41), 12663-12668. https://doi.org/10.1073/pnas.1419773112

Royal College of Speech and Language Therapists. (2012). *Speech, Language and Communication Needs in the Criminal Justice System and Best Practice responses to these.*

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68-78. https://doi.org/10.1037/0003-066X.55.1.68

Ryan, A., Gibbon, F. E., & O'shea, A. (2015). Expressive and receptive language skills in preschool children from a socially disadvantaged area. *International Journal of Speech-Language Pathology, 18*(1), 41–52. https://doi.org/10.3109/17549507.2015.1089935

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. Science, 274(5294), 1926-1928. https://doig.org/10.1126/science.274.5294.1926

Sandelowski, M. (2000). Combining qualitative and quantitative sampling, data collection, and analysis techniques in mixed-method studies. *Research in Nursing & Health, 23*(3), 246-255. https://doi.org/10.1002/1098-240X(200006)23:3<246::AID-NUR9>3.0.CO;2-H

Sandhofer, C. M., Smith, L. B., & Luo, J. (2000). Counting nouns and verbs in the input: Differential frequencies, different kinds of learning? *Journal of Child Language, 27*(3), 561-585. https://doi.org/10.1017/S0305000900004256

Schwab, J. F., & Lew-Williams, C. (2016). Language learning, socioeconomic status, and child-directed speech. *Cognitive Science, 7*(4), 264-275. https://doi.org/10.1002/wcs.1393

Schwartz, R. G., & Terrell, B. Y. (2008) The role of input frequency in lexical acquisition. *Journal of Child Language, 10*(1), 57-64. https://doi.org/10.1017/S0305000900005134

Sénéchal, M. (1997). The differential effect of storybook reading on preschoolers' acquisition of expressive and receptive vocabulary. *Journal of child language, 24*(1), 123-138. https://doi.org/10.1017/S0305000996003005

Sénéchal, M., & LeFevre, J. (2001). Storybook reading and parent teaching: links to language and literacy development. *New Directions for Child & Adolescent Development, 92,* 39-52. https://doi.org/10.1002/cd.14

Shanker, S. G., & Taylor, T. J. (2001). The House that Bruner Built. In S. G. Shanker, & T. J. Taylor (Eds.). *Jerome Bruner: Language, Culture and Self* (pp.50-70). Sage Publications Ltd.

Sharot, T., Korn, C. W., & Dolan, R. J. (2011). How unrealistic optimism is maintained in the face of reality. *Nature Neuroscience, 14*(11), 1475-1479. https://doi.org/10.1038/nn.2949

Shavit, Y., & Blossfeld, H. (1993). *Persistent Inequality: Changing Educational Attainment in Thirteen Countries. Social Inequality Series.* Westview Press.

Schneidman, L. A., & Goldin-Meadow, S. (2012). Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science, 15*(5), 659-673. https://doi.org/10.1111/j.1467-7687.2012.01168.x

Shotts, L. L., McDaniel, M., & Neeley, R. A. (2008). The impact of prolonged pacifier use on speech articulation: a preliminary investigation. *Contemporary Issues in Communication Science and Disorders, 35*(Spring), 72-75.

Shriberg, L. D. Flipsen, P., Thielke, H., Kwiatkowski, J., Kertoy, M. K., Katcher, M. L., Nellis, R. A., & Block, M. G. (2000). Risk for speech disorder associated with early recurrent otitis media with effusion: two retrospective studies. *Journal of Speech, Language, and Hearing Research, 43*(1), 79-99. https://doi.org/10.1044/jslhr.4301.79

Siddiqi, A., Subramanian, S.V., Berkman, L., Hertzman, C., & Kawachi, I. (2007). The welfare state as a context for children's development: a study of the effects of unemployment and unemployment protection on reading literacy scores. *International Journal of Social Welfare, 16*(4), 314-325. https://doi.org/10.1111/j.1468-2397.2007.00501.x

Siu, A. L. (2015). Screening for speech and language delay and disorders in children aged 5 years or younger: US Preventive Services Task Force recommendation statement. *Pediatrics, 136*(2), e474-e481. https://doi.org/10.1542/peds.2015-1711

Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.

Smith, G. (2018). Step away from stepwise. *Journal of Big Data, 5*(32), 1-12. https://doi.org/10.1186/s40537-018-0143-6

Snow, C. E., Pan, B. A., Imbens-Bailey, A., & Herman, J. (1996). Learning how to say what one means: A longitudinal study of children's speech act use. *Social Development, 5*(1), 56-84. https://doi.org/10.1111/j.1467-9507.1996.tb00072.x

Snowling, M. J., Hulme, C., Bailey, A. M., Stothard, S. E., & Lindsay, G. (2011). *Better Communication Research Project: Language and Literacy Attainment of Pupils during Early Years and through KS2: Does Teacher Assessment at Five Provide a Valid Measure of Children's Current and Future Educational Attainments? Research Report.* Department for Education. https://assets.publishing.service.gov.uk/media/5a7a2e97ed915d6d99f5dbe8/DFE-RR172a.pdf

Sparling, J., & Meunier, K. (2019). Abecedarian: An Early Childhood Education Approach that has a Rich History and a Vibrant Present. *International Journal of Early Childhood, 51,* 207–216. https://doi.org/10.1007/s13158-019-00247-2

Sparling, J., Ramey, S. L., & Ramey, C. T. (2021). Mental Health and Social Development Effects of the Abecedarian Approach. *International Journal of Environmental Research and Public Health, 18*(13), 1-10. https://doi.org/10.3390/ijerph18136997

Spencer, S., Clegg, J., Stackhouse, J., & Rush, R. (2017) Contribution of spoken language and socio-economic background to adolescents' educational achievement at age 16 years. *International Journal of Language & Communication Disorders, 52*(2), 184-196 https://doi.org/10.1111/1460-6984.12264

Sperry, D. E., Sperry, L. L., & Miller, P. J. (2019). Reexamining the verbal environments of children from different socioeconomic backgrounds. *Child Development, 90*(4), 1303-1318. https://doi.org/10.1111/cdev.13072

Spiegel, C., & Halberda, J. (2011). Rapid fast-mapping abilities in 2-year-olds. *Journal of Experimental Child Psychology, 109*(1), 132-140. https://doi.org/10.1016/j.jecp.2010.10.013

Stanford Encylopedia of Philosophy (n.d.). In *Stanford Encylopedia of Philosophy.* Retrieved July 31, 2024, from https://plato.stanford.edu/index.htmlhttps://dictionary.cambridge.org/dictionary/english/word

StatsWales. (2023). *Highest qualification levels of working age adults by qualification, UK country and year.* Welsh Government. https://statswales.gov.wales/Catalogue/Education-and-Skills/Post-16-Education-and-Training/Lifelong-Learning/Qualification-Levels/highestqualificationlevelsofworkingageadults-by-qualification-ukcountry-year

Stefanowitsch, A. (2020). *Corpus linguistics: A guide to the methodology*. Language Science Press. https://doi.org/10.5281/zenodo.3735822

Stokes, S. F. (2010). Neighborhood density and word frequency predict vocabulary size in toddlers. *Journal of Speech, Language, and Hearing Research, 53*(3), 670-683. https://doi.org/10.1044/1092-4388(2009/08-0254)

Storkel, H. L. (2004). Do children acquire dense neighborhoods? An investigation of similarity neighborhoods in lexical acquisition. *Applied Psycholinguistics, 25*(2), 201-221. https://doi.org/10.1017/S0142716404001109

Storkel, H. L., & Hoover, J. R. (2010). An online calculator to compute phonotactic probability and neighborhood density on the basis of child corpora of spoken American English. *Behavior Research Methods, 42*(2), 497-506. https://doi.org/10.3758/BRM.42.2.497

Storkel, H. L., Maekawa, J., & Hoover, J. R. (2010). Differentiating the Effects of Phonotactic Probability and Neighbourhood Density on Vocabulary Comprehension and Production: A Comparison of Preschool Children With Versus Without Phonological Delays. *Journal of Speech, Language, & Hearing Research, 53*(4), 933-949. https://doi.org/10.1044/1092-4388(2009/09-0075)

Stromswold, K. (2001). The heritability of language: A review and metaanalysis of twin, adoption, and linkage studies. *Language, 77*(4), 647-723.

Strutt, C., Khattab, G., & Willoughby, J. (2021). Does the duration and frequency of dummy (pacifier) use affect the development of speech? *International Journal of Language & Communication Disorders, 56*(3), 512-527. https://doi.org/10.1111/1460-6984.12605

Sunstein, C. R. (2022). The distributional effects of nudges. *Nature Human Behaviour, 6*(1), 9-10. https://doi.org/10.1038/s41562-021-01236-z

Swingley, D., & Aslin, R. N. (2007). Lexical competition in young children's word learning. *Cognitive Psychology, 54*(2), 99-132. https://doi.org/10.1016/j.cogpsych.2006.05.001

Swingley, D., & Humphrey, C. (2018). Quantitative linguistic predictors of infants' learning of specific English words. *Child Development, 89*(4), 1247-1267. https://doi.org/10.1111/cdev.12731

Szreter, S. (2021). *The history of inequality: the deep-acting ideological and institutional influences.* Institute for Fiscal Studies. https://ifs.org.uk/inequality/wp-content/uploads/2021/11/The-history_of-inequality-final-IFS-Deaton-Review.pdf

Tamis-LeMonda, C. S., Shannon, J. D., Cabrera, N. J., & Lamb, M. E. (2004). Fathers and mothers at play with their 2-and 3-year-olds: Contributions to language and cognitive development. *Child Development, 75*(6), 1806-1820. https://doi.org/10.1111/j.1467-8624.2004.00818.x

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Current Directions in Psychological Science, 23*(2), 121-126. https://doi.org/10.1177/0963721414522813

Tardif, T., Fletcher, P., Liang, W., Zhang, Z., Kaciroti, N., & Marchman, V. A. (2008). Baby's first 10 words. *Developmental Psychology, 44*(4), 929-938. https://doi.org/10.1037/0012-1649.44.4.929

TedX Talks. (2014, October 28). *How to motivate yourself to change your behaviour* [Video] Youtube. https://www.youtube.com/watch?v=xp0O2vi8DX4

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.

Theakston, A. L., Lieven, E. V., Pine, J. M., & Rowland, C. F. (2001). The role of performance limitations in the acquisition of verb-argument structure: An alternative account. *Journal of Child Language, 28*(1), 127-152. https://doi.org/10.1017/S0305000900004608

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: use of stress and statistical cues to word boundaries by 7-to 9-month-old infants. *Developmental Psychology, 39*(4), 706-716. https://doi.org/10.1037/0012-1649.39.4.706

Thomas, M. S., Van Duuren, M., Purser, H. R., Mareschal, D., Ansari, D., & Karmiloff-Smith, A. (2010). The development of metaphorical language comprehension in typical development and in Williams syndrome. *Journal of Experimental Child Psychology, 106*(2-3), 99-114. https://doi.org/10.1016/j.jecp.2009.12.007

Tindbaek, P. (2019, May 23). The Program Where Older Volunteers Become School 'Aunts' And 'Uncles'. *Forbes*. https://www.forbes.com/sites/nextavenue/2019/05/22/the-program-where-older-volunteers-become-school-aunts-and-uncles/?sh=41b20c631086

Tomasello, M. (2005). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

Tomasello, M., Conti-Ramsden, G., & Ewert, E. (1990). Young children's conversations with their mothers and fathers: differences in breakdown and repair. J*ournal of Child Language, 17,* 115-130. https://doi.org/10.1017/S0305000900013131

Tomasello, M., & Haberl, K. (2003). Understanding attention: 12-and 18-month-olds know what is new for other persons. *Developmental Psychology, 39*(5), 906-912. https://doi.org/10.1037/0012-1649.39.5.906

Tracey, L., Bowyer-Crane, C., Bonetti, S., Nielsen, D., D'Apice, K., & Compton, S. (2022). *The Impact of the COVID-19 Pandemic on Children's Socio-Emotional Wellbeing and Attainment during the Reception Year. Research Report*. Education Endowment Foundation. https://files.eric.ed.gov/fulltext/ED620337.pdf

Tracy-Ventura, N., Huensch, A., & Mitchell, R. (2021). Understanding the long-term evolution of L2 lexical diversity: The contribution of a longitudinal learner corpus. In B. Le Bruyn, & M. Paquot (Eds.), *Learner Corpus Research Meets Second Language Acquisition* (pp.148-171). Cambridge University Press.

Tribushinina, E. (2015). The Role of Paradigmatic Semantic Relations in Adjective Acquisition: Evidence from Two Russian-Speaking Children. In E. Tribushinina, M. Voeikova, & S. Noccettii (Eds.), *Semantics and Morphology of Early Adjectives in First Language Acquisition* (pp. 218−242). Cambridge Scholars Publishing.

Tribushinina, E., Van Den Bergh, H., Kilani-Schoch, M., Aksu-Koç, A., Dabašinskienė, I., Hrzica, G., Korecky-Kröll, K., Noccetti, S., & Dressler, W. (2013). The role of explicit contrast in adjective acquisition: A cross-

linguistic longitudinal study of adjective production in spontaneous child speech and parental input. *First Language, 33*(6), 594-616. https://doi.org/10.1177/01427237135031

Trickett, E., Rasmus, S. M., & Allen, J. (2019). Intervention fidelity in participatory research: a framework. *Educational Action Research, 28*(1), 128–141. https://doi-org.ntu.idm.oclc.org/10.1080/09650792.2019.1689833

Tucker-Drob, E. M., & Bates, T. C. (2016). Large cross-national differences in gene× socioeconomic status interaction on intelligence. *Psychological Science, 27*(2), 138-149. https://doi.org/10.1177/0956797615612727

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*(2), 207-232. https://doi.org/10.1016/0010-0285(73)90033-9

University of Melbourne (n.d.). *Abecedarian Approach Australia.* https://education.unimelb.edu.au/3a

Vaden, K.I., Halpin, H.R., Hickok, G.S. (2009). Irvine Phonotactic Online Dictionary, Version 2.0. [Data file]. Available from http://www.iphod.com.

Van Gestel, L. C., Adriaanse, M. A., & De Ridder, D. (2021). Do nudges make use of automatic processing? Unraveling the effects of a default nudge under type 1 and type 2 processing. *Comprehensive Results in Social Psychology, 5*(1-3), 4-24. https://doi.org/10.1080/23743603.2020.1808456

Van Heuven, W. J., Mandera, P., Keuleers, E., & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English. *Quarterly Journal of Experimental Psychology, 67*(6), 1176-1190. https://doi.org/10.1080/17470218.2013.850521

Venema, T. A., Kroese, F. M., Verplanken, B., & de Ridder, D. T. (2020). The (bitter) sweet taste of nudge effectiveness: The role of habits in a portion size nudge, a proof of concept study. *Appetite, 151*, 1-6. https://doi.org/10.1016/j.appet.2020.104699

Visser, I., Geanbasu, A., Bergmann, C., Byers-Heinlein, K., Doyle, F. L., Hannon, E. et al. (2021). ManyBabies 3: A multi-lab study of infant algebraic rule learning. PsyArXiv. https://doi.org/10.31234/osf.io/aex7v.

Vlach, H., & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children's retention of learned words. *Frontiers in Psychology, 3*(46), 1-8. https://doi.org/10.3389/fpsyg.2012.00046

Von Glaserfeld, E. (1993). Questions and answers about radical constructivism. In K. G. Tobin (Ed.), *The practice of constructivism in science education* (pp. 23-38). Routledge.

Waddington, C. H. (1957). *The strategy of the genes*. Routledge.

Waite, S. (2018). Embracing Audiobooks as an Effective Educational Tool. [Master's thesis: The College at Brockport: State University of New York]. Digital Commons @ Brockport. https://core.ac.uk/download/pdf/233578849.pdf

Walker, D., Sepulveda, S. J., Hoff, E., Rowe, M. L., Schwartz, I. S., Dale, P. S., Peterson, C. A., Diamond, K., Goldin-Meadow, S., Levine, S. C., Wasik ,B. H., Horm, D. M., & Bigelow, K. M. (2020). Language intervention research in early childhood care and education: A systematic survey of the literature. *Early Childhood Research Quarterly, 50*, 68-85. https://doi.org/10.1016/j.ecresq.2019.02.010

Walmsley, J., & Mannan, H. (2009). Parents as co-researchers: a participatory action research initiative involving parents of people with intellectual disabilities in Ireland. *British Journal of Learning Disabilities, 37*(4), 271-276. https://doi.org/10.1111/j.1468-3156.2009.00582.x

Wasik, B. A., Hindman, A. H., & Snell, E. K. (2016). Book reading and vocabulary development: A systematic review. *Early Childhood Research Quarterly, 37*, 39-57. https://doi.org/10.1016/j.ecresq.2016.04.003

Watkins, M. P., & Meredith, W. (1981). Spouse similarity in newlyweds with respect to specific cognitive abilities, socioeconomic status, and education. *Behavior Genetics, 11*, 1-21.

Weisleder, A., & Fernald, A. (2013). Talking to children matters: Early language experience strengthens processing and builds vocabulary. *Psychological Science, 24*(11), 2143-2152. https://doi.org/10.1177/0956797613488145

Weizman, Z. O., & Snow, C. E. (2001). Lexical output as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning. *Developmental Psychology, 37*(2), 265-279. https://doi.org/10.1037/0012-1649.37.2.265

Winter, R. (2002). Truth or fiction: problems of validity and authenticity in narratives of action research. *Educational Action Research, 10*(1), 143–154. https://doi.org/10.1080/09650790200200178

Wolfson, G. (2008). Using audiobooks to meet the needs of adolescent readers. *American Secondary Education, 36(2)*, 105-114.

Wood, C. (2002). Parent-child pre-school activities can affect the development of literacy skills. *Journal of Research in Reading, 25*(3), 241-258. https://doi.org/10.1111/1467-9817.00173

Wood, W., & Rünger, D. (2016). Psychology of habit. *Annual Review of Psychology, 67*, 289-314. https://doi.org/10.1146/annurev-psych-122414-033417

Woodward, A. L., & Markman, E. M. (1991). Constraints on learning as default assumptions: Comments on Merriman and Bowman's "The mutual exclusivity bias in children's word learning.". *Developmental Review, 11*(2), 137-163.

Xin, R., & Brion-Meisels, G. (2022). Teachers' experiences with agency and well-being during a critical participatory action research project. *Educational Action Research, 32*(1), 72–89. https://doi-org.ntu.idm.oclc.org/10.1080/09650792.2022.2058973

Yeomans-Maldonado, G., Justice, L. M., & Logan, J. A. (2019). The mediating role of classroom quality on peer effects and language gain in pre-kindergarten ECSE classrooms. *Applied Developmental Science, 23*(1), 90-103. https://doi.org/10.1080/10888691.2017.1321484

Zemencuk, J., Rogosch, F. A., & Mowbray, C. T. (1995). The seriously mentally ill woman in the role of parent: Characteristics, parenting sensitivity, and needs. *Psychosocial Rehabilitation Journal, 18*(3), 77-92. https://doi.org/10.1037/h0095499

Zientarksi, D.P. & Pottorff, D.D. (1994). Reading Aloud to Low Achieving Secondary Students. *Reading Horizons, 35*(1), 44-51.