# Relating Language Input to Language Processes Early in Development

## Using the Early Language Processing Task in UK and India

Laia Fibla Reixachs

100219137

School of Psychology

University of East Anglia

A thesis submitted for the degree of

*Doctor of Philosophy*

June 2021

# Abstract

Early language development is highly associated with language outcomes and later cognitive abilities. This is why it is crucial to understand the mechanisms that support children's language acquisition, and the factors that influence it. To date, most studies use indirect measures of language skills. There are a few studies that use direct measures of the processes that support word learning, but they usually focus only on one particular mechanism. Thus, we do not know how these early language processes relate to one another. Moreover, research has shown that a critical factor in early language development is the linguistic environment that children grow up in, particularly, the language input that they are exposed to. However, not many studies relate children's language environment with their later linguistic abilities, and only a handful do that longitudinally. The present thesis aims to contribute to this body of research by developing an early language task that includes several measures of language processing. This new task is used to measure the relationship between early language input, vocabulary knowledge and language processing at different time points from infancy to toddlerhood. Moreover, it extends this research to an at-risk population of Indian children. Study 1 (Chapter 2) shows the development of the Early Language Processing (ELP) task, a direct measure of language processing that is able to capture different language processes in a sample of UK children. This study includes data from the same sample of children at two time points; when they were 15 to 27-months-old (test), and when they were 28 to 36-months-old (retest). Results show strong developmental effects, as well as individual differences, replicating findings from previous literature. We also find positive relationships between ELP and other well established measures of vocabulary knowledge. In Study 2 (Chapter 3), we investigate the relationship between early language input in both infancy and toddlerhood (6 and 18 months of age), and language processing abilities measured with ELP. We find relationships between language input and language processing that suggest that children might benefit from different aspects of input at different ages. Furthermore, children seem to benefit from amount of adult words and from conversational experience in different ways, depending on the process measured by ELP. In Study 3 (Chapter 4), we administer the ELP task to an at-risk sample of children based in India varying in socio-economic status (SES), from similar ages to our sample collected in the UK. We successfully translate and adapt the task to a different language and population. Our findings show that only older Indian children from higher SES show similar looking patterns to those seen in the UK sample. Study 4 (Chapter 5) relates Indian children language input measures collected in their homes, to their ELP measures. Our results show different relationships between language input and language processes across ages and SES, and highlight the importance of collecting data from understudied populations using multiple measures. Overall, the data presented in this thesis contributes to research in language development by creating a new early measure of language processing abilities, and using it to relate children's early language experiences with their language skills. Importantly, we expand this research to a new population in rural India, adding data from understudied settings to our knowledge of the processes that support word learning. Our results set the stage for future work to measure how early language processes predict long-term language and cognitive abilities.

# Acknowledgements

*Gràcies*

I want to thank my family, both Fibla and Reixachs. En particular el Joan, l'Àngels, el Jan i el Manel per sempre estar al meu costat i recolzar-me en cada pas del camí, tot i que de vegades aquest camí em duu lluny de casa. Vosaltres sou l'energia, la força i el l'afecte que guia cada pas que faig cap a nous horitzons. A vosaltres, us dedico aquesta tesi. En especial al meu pare Joan, a qui li feia molta il·lusió veure el fruit de tots aquests anys de feina. Encara que ja no sigui aquí físicament per veure-ho, sempre m'acompanya.

I also want to thank all the inspirational woman that I have met along the way. From well known researchers, to the most humble ladies. To all of you, thank you for each piece of knowledge that you have passed to me. Les més importants inclouen la meva mare, les meves àvies Teresa i Nuri, i la meva besàvia Rosa. Una part d'aquestes dones també inclou Les Fantàstiques: Cris, Anna, Ire, Laura, Mariana, Nerea, Laia, Joana, Júlia, Maia i Jana. A les que dono les gràcies per compartir les erniquidores històries del seu dia a dia amb mi.

I also want to thank all the friends I made during my journey from Paris, to Norwich and Barcelona. Particularly Sonja, Camila and Therese for always being there for me like true sisters. I want to thank all the members of the Swing dance community in Norwich, which have been a crucial part of my life during my PhD years. Thank you for your friendship, all the laughs, good times and, of course, dances. I would like to particularly mention Julia, Andy, Hannah, Ali and Dave, Helga, Leo and Lucy, Jess and Phil, Lucy, Ollie and Sally-Anne. For their invaluable friendship and support during good times, as well as bad times.

This project would not have been possible without all the children and families from UK and India that gave their time to science, and participated in the studies of this thesis. Simply by looking or speaking, each one of them contributes to being a tiny bit closer to understanding the complexity of language development. This project would neither exist without the help, guidance, support and dedication of my mentors, Larissa Samuelson and John Spencer. I particularly would like to thank them for the freedom they gave me while designing this project at the beginning of my PhD, and for fully trusting me to carry it on in rural India, even though the adversities. Over the last years, they have treated me like an academic daughter by always supporting and enhancing my ideas, which has helped me develop multiple skills that I consider essential for infant research. Moreover, they have also taught me that it is possible to combine pursuing an academic career while having a family. For all that, you will always have a special place in my heart.

A special thank you goes to my co-workers Sam, Prerna, Lourdes, Laura, Milena and Ellie. For their advice, guidance while coding my data, laughs, stats tricks, coffees and dinners shared, and for their support throughout this project. I would also like to thank the other PhD's from my cohort at UEA, particularly Theodora, Carys and Sara.

I would like to acknowledge the members of the Developmental Dynamics Lab at UEA who have helped facilitate my projects throughout my PhD. In particular, I would like to thank the RAs who have contributed to my projects:

Rosie, Jess and Giulia. Also, thank you to the data collection assistants whose work was invaluable to the completion of this project, especially Rachel, Sehar, Nicky, Kate and Kiara. Rachel was particularly kind to let me use her voice for the stimuli in English. Thank you to the students that volunteered to help in different tasks related to this project, in particular Natali, Izzy, Chloe and Anjelee. Additionally, I would like to thank Jordan for his help coordinating the LENA data collection. As well as Jo and Laura, the DDLab's project participant coordinators, for *everything* they do. Jo deserves a special mention for always being *a star*.

Extending this project to India would not have been possible without the close collaboration between our laboratory and the Community Empowerment Lab. Each member from their field team in Shivgarh helped me in different ways in every step of this study, guiding me through a new society and culture. In particularly, I would like to thank Babita for lending her voice for the stimuli in Awadhi. Vineeta and Pooja for her dedication testing children on our task. Reetu for her good mood and being a bridge between me and the local families and children while testing. Jyoti and Jay for their help adapting the images to the local culture and translating the sentences from our task to Awadhi. Abhishek, Ankit and Vinay for their technical support and management of the project in India. Working with all this team in Shivgarh has been one of the highlights of my PhD, and an incredible life experience.

Some ideas for this project started back when I was a master student at the Laboratoire de Sciences Cognitives et Psycholinguistique in Paris. My years there working with Anne Christophe and Alejandrina Cristia were decisive in my interest to research language acquisition in relation to children's home experiences. Thus, I thank them, as well as all the other members of the LSCP for inspiring this project.

The last thank you is the deepest one. I would like to thank Keijo *mein schatz*, for being a real treasure and a real life companion over the years. Thank you for everything you do. This section needs a special mention to Marie, that has closely watched over the writing of this thesis from top to bottom.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Language is a unique capacity of humans, which we use to communicate with one another. Other species, such as dolphins, chimpanzees or even bees and ants, also have communicative abilities but no communicative system is as complex and sophisticated as human language.

Human language acquisition is not an easy task, it requires learning the meaningful sounds of a language, learning how to produce those sounds, learning how to segment the speech stream, learning the meaning of words, acquiring syntax, acquiring morphology, and using all those skills to communicate about a wide variety of topics in very different situations. Children exposed to more than one language develop this set of skills for each of the languages that they are learning. Indeed, human children across different languages and cultures learn their language (or languages) very quickly, in a process that might seem effortless. The question is, then, how do children do this?

Focusing only on the apparently simple task of learning a new word, we see the complexity of acquiring a language. Imagine a sunny day of spring. A mother and her toddler daughter, Julia, are playing with some wooden building blocks in a garden full of flowers and blossoming trees. Suddenly, a gray cat jumps from the neighbour's fence into the garden. Mum says, "Look, a kitty!". We might think that in that instance Julia would learn that *kitty* means *cat*. Unfortunately, the word learning task is usually not that simple. First, Julia needs to extract the word *kitty* from continuous running speech, which has no apparent boundaries between words. Research has shown that children are able to use different cues found in the structural regularities of a language to distinguish boundaries between words and extract words from speech; these cues include prosodic patterns, phonotactic regularities, and distributional properties of words (Saffran, Aslin, & Newport, 1996; Jusczyk, 1999).

Even if Julia is able to extract *kitty* from that sentence, how does Julia know what mum is referring to if she has never heard the word *kitty* before? It could mean flower, tree, a specific type of flower or tree, animal, the color of the cat, the action of jumping, among many other possibilities. A possible solution that would help Julia associate the new word with the correct object would be to look for the most novel item. However, the blossoming garden is full of things that are new to her, so how does she know which one mum was referring to? Moreover, she might have been looking at the wooden building blocks when mum said the word *kitty*, because they were novel and thus very exiting. This would complicate the word learning scenario even more, because Julia was paying attention to the wrong novel object at the time the new word was said. This attraction to novelty is known as novelty bias (i.e., the attraction of young children towards novel objects), and it could help children to find the correct object for a new word in some situations. In other situations, children might need to look for additional cues beyond novelty to find the correct referent. For instance, it is possible that Julia might already know the meaning of the word *cat*, which would imply having to associate two words with the same referent. This is often the case for children growing up in multilingual contexts who must solve the reference problem known as ”the problem of the radical translation” (Quine, 1960), the idea that a new word could mean many different things. How can Julia distinguish between all the infinite possible meanings of a word, given that she does not know much about language?

Research suggests that Julia has some strategies that help her to determine the likely meaning of the new word *kitty*. For example, her vocabulary skills – the previous knowledge of the meaning of some words – could help her reduce the referents of the new word. However, in this case, this might not be enough because, as we can see in this example, some referents can have two labels (i.e., cat and kitty). Luckily for her, it is possible that mum pointed at the cat while uttering the sentence, which could help Julia refocus her attention from the wooden building blocks to the cat, the correct referent. However, even if Julia was able to map the correct word into the correct referent, she needs to remember this association and extend it to all *kitty* categories (i.e., cats from different colours, sizes, even cartoons cats) in order to continue building her vocabulary and her overall language skills.

As we can see, learning words is not easy. In the word learning task, children need to master a set of skills and use them to learn new words successfully. Those skills include language processes and cognitive abilities which children might combine in different ways, based on their previous experiences, their age, their individual abilities, and the context in which word learning occurs. This set of skills include comprehension and production vocabulary which could help discard some of the possible referents, word processing abilities that help children quickly identify a word after hearing it, the ability to inhibit attraction to novelty when other cues indicate that something else is the correct referent, and the capacity to remember word-object mappings later on to continue building vocabulary.

At this point word learning might seem a really difficult task, however, we know that children are quite good at it. At 6 months of age, children learning American English already know the meaning of some common words such as body parts, foods or ”mommy” (Bergelson & Swingley, 2012). Parental report data shows that 12-month-old children are able to comprehend 77 words on average, with the lower 10 percent of children understanding 21.2 words (i.e., decile 0.10) and the highest 10 percent of children understanding 185 words (i.e., decile 0.90). The average of words that children are able to understand increases dramatically by 18 months of age, with children being able to comprehend 244 words on average, with the lower 10 percent of children understanding 112.5 words (i.e., decile 0.10) and the highest 10 percent of children understanding 356.9 words (i.e., decile 0.90). This means that in only six months, children with typical development have acquired a comprehension vocabulary of almost 100 words more on average (examples from data retrieved from the Wordbank database based on MCDI questionnaires; Frank, Braginsky, Yurovsky, & Marchman, 2017). However, the variability seen in this data is striking. Twelve-month-olds with slower vocabulary development comprehend only 21.2 words, whereas those with the fastest development understand up to 185 words – notably more than the slower 18-month-old children who comprehend only 112.5 words. This same variability in word comprehension scores can be seen in other languages such as Croatian, Danish, French, Hebrew, Italian, Kigiriama, Korean, Mandarin, and Russian (see Wordbank database for examples; Frank et al., 2017). It is clear from this example that children vary substantially in early lexical development.

The same amount of variability is found when looking at word production.

Over the second year of life, some children show rapid growth or vocabulary production speaking more than 250 words by the age of 18 months, while others do it more slowly, speaking fewer than 10 words at this age (Fenson, Marchman, Thal, Dale, & Reznick, 2007). Although delayed onset of expressive language can be a risk factor, potentially leading to later language and academic difficulties (Rescorla, 2009), usually these early delays are not critical for most children because nearly two thirds of late-talkers move into the normal range before preschool. The remaining third however, will have persistent language difficulties at 3 and 4 years of age (Dale, Price, Bishop, & Plomin, 2003).

Possible explanations of these differences in vocabulary development might lie in children's early language processing abilities and the early language environment. Literature has examined which variables predict late talker's catch up. A study by A. Fernald and Marchman (2012) showed that late talkers who were more efficient in word recognition at 18 months were also more likely to "bloom," showing more accelerated vocabulary growth over the following year, compared to late talkers who were less efficient in early speech processing. This shows that there might be robust links between processing efficiency and vocabulary growth during toddlerhood in late talking children at 18 months. Other studies have focused on contextual variables, reporting that late talkers from lower socioeconomic status (SES) families are more likely to have persistent language difficulties long term (Rescorla, 2011; Armstrong et al., 2017). In a meta-analysis, significant predictors of expressive-language outcomes in later talkers included toddlerhood expressive-vocabulary size, receptive language, and SES (Fisher, 2017). These findings indicate that differences in children's language processing efficiency, expressive and receptive vocabulary, as well as SES have cascading consequences for later learning and may contribute to the individual differences in language proficiency observed across children. Moreover, variation in early language skills has also been documented across other language abilities such as speech perception, segmentation, and recognition skills, and those early differences are also predictive of children's vocabulary measures in toddlerhood (Cristia, Seidl, Junge, Soderstrom, & Hagoort, 2014). Thus, it is important to understand the sources of this variation and how they affect the different processes involved in word learning. The previous studies also highlight the role of children's environment, such as their SES, in children's language abilities and long term language outcomes. Therefore, we need to study both early word learning and also take into account the context where word learning occurs.

Many prior studies have explained individual differences in early word learning based on children's particular experiences with language (Hoff, 2006, for a review). The two main contextual factors that have been associated with variation in children's language skills are the amount of language exposure (or language input) that the child receives at home, and the SES of the family using indices such as parental education (from both or only the primary caregiver) and/or income. For instance, children from lower-SES households receive less language input than their higher-SES peers, and that quantity and quality of parental input is associated with children's rate of vocabulary growth (e.g., Hart & Risley, 1992, 1995; Rowe, 2012). Thus, higher-SES infants who are exposed to larger quantities of parental speech and a richer language input show better vocabulary skills later on. Studies with children learning more than one language also show that differences between individual bilingual children's use of their two languages can be attributed to differences in the language input environments for each of the languages. Those differences include the child's age of first regular exposure to each language, relative and absolute frequencies of input for each language, or parent's interaction strategies using those languages (e.g., Houwer, 2011). Amount of input is also predictive of children's processing abilities. Children exposed to more maternal speech early in development know more words and are faster at word recognition later on (Hurtado, Marchman, & Fernald, 2008a). It is likely that the effect of language input on children's vocabulary is based on infants' language-processing efficiency, because richer language experiences help children's processing skills which facilitate language growth (Weisleder & Fernald, 2013).

The previous literature shows that to be able to understand vocabulary development, we need to understand the multiple processes involved in word learning and how differences in the amount of input to children influence vocabulary learning. However, measuring early word-learning skills is not easy in pre-verbal infants. This is why most vocabulary measures are indirect, based on parental report. Vocabulary checklists can be very useful for tracking children's attainment of standard language milestones, and can be used to compare individual and group data on language development, but parental report – particularly of word comprehension – can be very hard. It is possible that parents underestimate or overestimate their child's comprehension abilities. Moreover, parents' criteria for what constitutes a 'known' word might change as the child gets older (see Tomasello & Mervis, 1994 commentary on Fenson et al., 1994 monograph). An alternative is to use measures of vocabulary development or early grammar (e.g., Mullen scales of early learning Mullen & Others, 1995), but these are generally indirect, focusing on higher-level outcomes and not on underlying language processes. In addition, some of these outcome measures are only suitable for older children.

There are some direct early measures of language skill, such as the looking while listening paradigm (A. Fernald, Zangl, Portillo, & Marchman, 2008). While useful early, such measures typically focus only on a single aspect of word learning (e.g., speed of language processing) rather than integrating across multiple measures relevant for the word learning task. This makes it hard to know how the multiple processes involved in word learning relate to one another. Vocabulary development is the result of several language processing skills that children use to learn a word. To be able to understand lexical development, we need to understand how those processes work together, and how they influence word learning.

A further problem is that most of the literature on early word learning comes from western societies. This makes it hard to generalize findings across cultures and populations. Extending early language research to children growing up in different cultural contexts is crucial given the important role that the environment plays in children's language development. The gap of studies from non-western contexts might be related to the lack of tasks that can be used with young infants in cross-cultural contexts. Thus, it would be ideal to develop tasks that can be used across multiple cultural environments. Finally, not many studies have related contextual factors such as language input and demographic information with multiple measures of language processing abilities in the same group of children over development. Even fewer studies have done this cross-culturally. Measuring multiple language processes over development across different populations could clarify why we see such variation in children's language skills during the first years of life and the role that the environment plays in children's language processing abilities.

The present project aims to contribute to our understanding of the relationships between early language processing abilities and language input in early infancy in both the UK and India. To measure children's language processing abilities, we developed the Early Language Processing (ELP) task, a looking while listening task that is able to directly and efficiently measure multiple language processes in individual children. The ELP task is based on several well established tasks that have been shown to be predictive of later language skill: speed of language processing, online word comprehension, novelty biases, referent selection, and retention of new words. We administered the ELP task to UK and and Indian children at multiple time points in early development (at approximately 18- and 30-months-of-age). To measure the relationships between language input and language processing, we gathered naturalistic recordings of children's language input at home during infancy and toddlerhood, and looked for relationships between language input and processing abilities. We also looked at how SES affected these relationships between input and language in our Indian sample.

## 1.1 Early Language Processes

Better understanding of the multiple processes involved in early word learning and language development requires measuring these processes in a way that allows examination of their relationships and how they change over development. The ELP task was developed with this aim in mind. It is the first measure to integrate several language processes together: speed of language processing, online word comprehension, novelty biases, referent selection, and retention of new words. These five measures are particularly interesting because studies have shown that they are predictive of later language outcomes and/or essential for children to learn a word. We detail them below and highlight work across different populations.

### Speed of Language processing and Comprehension Vocabulary

We have already mentioned the role of speed of word processing on language development. Both speed of word processing (SoP) and vocabulary early in development are predictive of later language and cognitive abilities. Speed of word processing measures how fast a child can recognise a spoken word and it is an indicator of how well a child knows a word, since children look faster to the images of the words that they know very well after hearing them (A. Fernald, Pinto, Swingley, Weinbergy, & McRoberts, 1998). Studies show that word processing speed is positively related to children's vocabulary size; children with faster speeds of word recognition have larger vocabulary sizes. For example, SoP measured as reaction time at 25 months is strongly related to children's vocabulary growth over the second year of life (A. Fernald, Perfors, & Marchman, 2006). Furthermore, children's speed of spoken word recognition and vocabulary size at 25 months are both predictive of later linguistic and cognitive skills at 8 years of age (Marchman & Fernald, 2008). Vocabulary in infancy alone has been shown to also be predictive of later language skills. For example, a study showed that vocabulary knowledge measured using parental report between 16 and 24 months was predictive of later vocabulary, phonological awareness, reading accuracy, and reading comprehension when children were between 4 and 9 years of age (Duff, Reen, Plunkett, & Nation, 2015). Thus, it is possible that both SoP and vocabulary in infancy work together as a platform for developing later language and cognitive skills, including literacy skills. This makes these two measures good candidates to include in the ELP task. However, to understand how they relate to word learning, we need to integrate them with basic word learning processes such as referent selection, retention of new words, and the possible biases that might influence word-referent associations, such as attraction to novelty.

## Novelty Biases, Referent Selection and Retention

The literature on speed of word processing suggests that children who rapidly recognize and interpret familiar words typically have accelerated lexical growth. This provides indirect evidence that lexical processing efficiency is related to word-learning ability, that is, the ability to map new words to new objects. In a recent study, Lany (2018) found a relationship between speed of lexical processing and novel word learning in 18-month-olds and 30-month-olds. Children who were faster at recognizing familiar words were also more accurate at recognizing novel words in a word learning task. This is evidence that in the task of learning a word, it is crucial to be able to quickly recognise words to build vocabulary skill, and that this helps when making new word-object associations. Another skill that children use to make new word-object associations is to pay attention to the most novel item. Novelty biases help children map novel names to novel referents rather than to familiar ones (Mather, 2013). In this process, children use prior lexical knowledge to determine the referent of a novel word via mutual exclusivity (Markman & Wachtel, 1988). Thus, children are able to map novel words onto novel objects in the context of familiar ones, if they know the label for the familiar object. Mutual exclusivity has been demonstrated in children from 14 to 30 months using multiple paradigms including 2-dimensional images on a screen (Bion, Borovsky, & Fernald, 2013) and 3-dimensional objects (Horst & Samuelson, 2008; Samuelson, Kucker, & Spencer, 2017).

Using 3-dimensional objects, the interaction between novelty driven attention and lexical knowledge has been evaluated in referent selection and retention tasks (RSR), in which children learn a new word using mutual exclusivity and then they are tested to see if they remember the new-word object association. These studies show the complexity of the role of novelty bias in word learning in two ways. First, children's attention to novelty changes over development and, second, even though attraction to novelty might be very useful when learning a new word-object mapping during referent selection, too much attraction to novelty can prevent children from retaining the new object-word association, particularly at younger ages (Kucker, McMurray, & Samuelson, 2018). Usually, too much attention to novelty occurs in cases where the knowledge of the familiar word is weak. Children need to have a good knowledge of the familiar word to use mutual exclusivity efficiently during referent selection – they can learn that the new label refers to the novel object by discarding that it does not refer to the familiar one. The effect of lexical knowledge in relation to novelty can be seen across development. Studies show that 18-month-old children, who are less experienced with language, have very strong novelty biases that prevent them from selecting the correct object during referent selection tasks because they consistently select the novel object with both known and novel names (Kucker et al., 2018). However, 24 month-old children, who have larger vocabularies, can overcome the novelty bias and correctly select a novel referent in response to a novel word. However, these children are not able to remember the novel word-object association after a 5 minute delay (Horst & Samuelson, 2008; Bion et al., 2013; Kucker & Samuelson, 2012). By 30-months of age, however, children were able to overcome the novelty bias to select the correct referent and remember novel name–referent mappings over a time delay (Bion et al., 2013; Spiegel & Halberda, 2011). Those older children have a good knowledge of the familiar item and more experience with language, and therefore they can effectively use mutual exclusivity during referent selection to map a new word to a novel object, overcome the novelty biases, and learn the association, remembering it after a short delay.

The relationships between mutual exclusivity and lexical knowledge have also been demonstrated in work looking at children's semantic networks (i.e., the links connecting related words in children's vocabularies). Studies have shown that semantic network structure is related to children's language learning biases such as mutual exclusivity. In one study, 24 months old children who had lexical networks with more connections were better at mutual exclusivity (Yurovsky, Bion, Smith, & Fernald, 2012). Moreover, semantic network structure usually reflects linguistic input structure (e.g., Hills, Maouene, Maouene, Sheya, & Smith, 2009; Amatuni & Bergelson, 2017). Studies measuring mutual exclusivity in multilingual children, which have multiple one-to-one word-meaning mappings and hear less language input in each language, show that vocabulary size is related to mutual exclusivity performance (Lewis, Cristiano, Lake, Kwan, & Frank, 2020). Linguistic experience in multilingual contexts shapes the development of mutual exclusivity use because bilingual children use social-pragmatic cues in addition to mutual exclusivity to learn new words (Kalashnikova, Mattock, & Monaghan, 2015). Furthermore, children exposed to three languages use mutual exclusivity to a lesser extent than bilingual children who also use mutual exclusivity more flexibly and less often than monolingual children (Byers-Heinlein & Werker, 2009). These studies highlight the role of lexical knowledge in word learning strategies such as mutual exclusivity as well as the influence of the child's linguistic environment, particularly the amount of input received, on the language processes involved in word learning.

The number of languages a child is exposed to shapes its language development but the type of language matters too. Some studies show that there are cross-linguistic differences in the onset of word comprehension and thus, not all languages are acquired at the same time. For example, Norwegian infants show first evidence of word comprehension at eight to nine months of age—rather than from six to seven months of age for English-learning infants (Kartushina & Mayor, 2019). Moreover, after children start acquiring words, English infants tend to show a noun bias because they usually learn nouns before any other word type (Goodman, Dale, & Li, 2008; Braginsky, Yurovsky, Marchman, & Frank, 2019), but in other languages such as Mandarin it appears that verbs and nouns are acquired more equally (Tardif, Gelman, & Xu, 1999). This has been explained due to the frequency of appearance of those words in a language (e.g., Roy, Frank, DeCamp, Miller, & Roy, 2015). In particular, because verbs are as frequent as nouns in Mandarin, children learn them at a similar rates.

There are also cross-linguistic differences in the acquisition of spacial semantic categories. For example, English and Korean differ in how they describe the location of an object in relation to other objects. English children learn that they have to use *on* for objects that are on a surface (i.e., support relations) and *in* for objects that are contained inside an enclosure (i.e., containment relations). Korean children do not have to do that because Korean does not distinguish between containment and support. Instead, Korean uses the verb *kkita* to describe objects that are tightly fitted into another object in an interlocking manner and a range of verbs to describe loose-fitting arrangement. Thus, Korean children will learn the verb *kkita* for ”putting a ring (on) a finger” or ”putting a cassette (in) a case” (Bowerman & Choi, 2001; including examples). These studies show examples of how children need to learn different things based on the language or languages that they are exposed to, which will shape their language skills and their overall language development.

Overall, the literature measuring early language processes shows the importance of lexical knowledge in relation to novelty biases, referent selection, and retention of new words as well as the association between speed of word processing, vocabulary abilities, and word learning. All these processes are part of the "big" task of learning a new word; thus, conceptually, it is expected that they influence each other. To date, however, no study has measured these processes in a single task. Moreover, studies measuring language processes early in development across diverse populations learning different languages are scarce, and we still do not know enough about how different environments shape early language processes. The present project aims to contribute to the field, documenting early language processes in different cultural contexts by adapting the ELP task to be used in rural India with children learning the Awadhi dialect. Gathering data that is representative of all learning experiences is important because those early experiences will influence how infants use their multiple language processing abilities in the word learning task. This could help clarify why there is large individual variation in early language development.

## 1.2 Early Language Experiences in relation to Language Processes

We have already mentioned the role of children's early environments in their language development, but what exactly influences emerging language abilities? The short answer is that language input plays a big role (for a review see Rowe & Weisleder, 2020). We review research on language input and environmental variables in the following, highlighting work across different cultures.

There is a large body of research showing how the amount of caregiver language input such as the number of words (Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991), the quality of that input such as the lexical diversity or sentence complexity (Rowe & Snow, 2020), conversational experience (Zimmerman et al., 2009), and parents use of language such as child directed speech (Rowe, 2012) is related to children's vocabulary size and later language abilities. Most of this literature reports differences in vocabulary size based on children's language experience. However, an important question is whether the beneficial effects of language input on later vocabulary size are constrained by the child's ability to efficiently process that input. Some studies have shown that maternal speech at 18-19 months is related to children's speech processing efficiency and vocabulary size at 18 and 24 months (Hurtado et al., 2008a; Weisleder & Fernald, 2013). A study found similar effects in older children with language input and lexical processing at 28–39 months predicting vocabulary size one year later in preschoolers (Mahr & Edwards, 2018). This literature suggests that larger quantities of language input provide children with more opportunities to practice recognizing words, which leads to greater processing efficiency, and facilitates word learning across development. Thus, language input is beneficial for language development because it contributes to children's language processing abilities as well as their language outcomes. However, children's linguistic environments vary across populations. Some children live in large families with many other children and, thus, they are exposed to adult input but also child input. Other children might spend most of their time with an adult caregiver, such as a nanny. Therefore, their linguistic experiences are mostly with a single adult. Children's linguistic environment might also be affected by cultural practices and beliefs about child development and how to talk to children, which might influence how parents talk to their children.

The role of child directed speech (CDS), a physically exaggerated and tonally high-pitched style of speech that adults use when talking to babies and young children, also know colloquially as ”baby talk” or ”motherese” (e.g., Snow, 1972), has had particular relevance in cross-cultural studies (see Lieven, 1994). In western populations, CDS has been reported to facilitate early word learning (Cartmill et al., 2013; Hoff, 2003; Rowe, 2008; Weisleder & Fernald, 2013) and it has been documented across a variety of languages (e.g., A. Fernald et al., 1989). However, in some cultures it has been reported to be almost non-existent. This is the case in a black community in Piedmont Carolinas (Heath, 1982, 1983), the Kaluli people of New Guinea, the American Samoa (Ochs, Schieffelin, et al., 1984) and the Javanese (Smith-Hefner, 1988). Recent studies in other Indigenous populations using modern techniques have also documented lower rates of CDS in those communities than those reported in western contexts. Casillas, Brown, and Levinson (2020) gathered daylong at-home audio recordings from children between 2 months and 3 years in a Tseltal Mayan village located in Southern Mexico. Using those recordings, the researchers measured how often children were engaged in verbal interaction with others and how the speech environment changed with age and context (e.g., household size, number of speakers present, or time of day). Children in this population were directly spoken to infrequently, with most directed speech coming from adults, and no increase with age.

Another study also showed that CDS is infrequent in the Tsimane community, a forager-farmer population (Cristia, Dupoux, Gurven, & Stieglitz, 2019). This means that for these children to learn language, they need to extract the information they need despite minimal speech from adults directed to them. This experience is drastically different form what have been reported in western cultures, where CDS leads to benefits in language outcomes. This suggests that children exposed to very little CDS might use other cues from adult speech to learn language that are unrelated to those characteristic of

CDS. Alternatively, children could still benefit from small amounts of CDS in those contexts. A study carried on in Yucatec Mayan children whose parents rarely engage in conversation with them – and, therefore, are exposed to very little CDS – showed that 2-year-old children's vocabulary size benefited from the number of different words that adults directed to them, even if that was a very small amount (Shneidman & Goldin-Meadow, 2012). Interestingly, only adult words, but not the number of words they were exposed to through overheard speech or speech directed to them by other children, were predictive of children's vocabulary size. This study shows that the amount of CDS is predictive of language development, even in cultures where adults direct very little speech to children and most of their input comes from overheard speech or speech directed from other children.

A key question is whether very little CDS leads to slower rates of language acquisition or developmental delays long term. Some evidence of the long term effects of low rates of CDS comes from a later study in the Tsimane community measuring phonological processing in children and adults. Phonological development is associated with lexical development (e.g., Werker & Curtin, 2005) and it has been related to language input experiences and literacy. Thus, it is particularly interesting to study phonological processing in populations with low rates of CDS and variation in literacy skills (Cristia, Farabolini, Scaff, Havron, & Stieglitz, 2020). This study found lower performance scores among Tsimane children and adults on a non-word repetition group game than those found in previous work with Italian speaking children (Piazzalunga, Previtali, Pozzoli, Scarponi, & Schindler, 2019) and Slovak speaking children (Kapalková, Polišenská, & Vicenová, 2013). This suggests that low levels of CDS in infancy has long-term effects on phonological processing. Since phonological development is associated with lexical development, the effect of low rates of CDS could also be present in lexical processes. Given the small amount of evidence we have regarding language development in communities similar to the Tsimane, more work is needed to determine the the short and long term impact of low rates of CDS characteristic of those societies. However, western studies have broadly documented that differences in parental speech addressed to the child based on SES are related to differences in children's language development.

Differences in the amount and quality of parental speech in western contexts have often been associated to the family SES. This is because many studies report that exposure to language input differs across social classes. These SES differences have been shown to explain variation in young children's vocabulary skills (e.g., Hart & Risley, 1992, 1995; Hoff, 2003). Children from higher-SES families hear more words and a richer input showing more vocabulary growth than children growing up in lower-SES families. Higher-SES children do not only benefit from hearing many words but their parents also respond more to them, produce more affirmative and encouraging utterances and fewer prohibitions (Hart & Risley, 1992), use more diverse words, longer sentences (Huttenlocher, Waterfall, Vasilyeva, Vevea, & Hedges, 2010; Rowe, 2008) and more wh-questions, which benefits children's language abilities and later child language at school age (Vernon-Feagans, Bratsch-Hines, Reynolds, & Willoughby, 2020).

Since family SES predicts children's language skills, some studies have investigated how SES exerts this influence and why we find lower quantity and quality of language input in lower SES families. A proposal is that SES affects early vocabulary development via maternal speech which differs based on maternal education (Hoff, 2003). Mothers or parents with higher education provide more and richer language-learning experiences which are beneficial for language development. This could be for three reasons (which are not necessarily mutually exclusive):

*Reason 1)* Parents with higher levels of education have better language abilities and, thus, provide richer input (Street and Dabrowska, 2010).

*Reason 2)* SES differences might reflect differences in language use based on beliefs about the value of talking to children and ideas about children development in general (Hoff, 2003; Rowe, 2008). Cultural beliefs around the world affect how parents talk and interact with their children. In some cultures, parents rarely engage with their children in linguistic interactions because it is generally believed that children have no understanding of the world. This has been reported to be the case among the Kaluli of Papua New Guinea (Ochs et al., 1984) and the K'iche' Mayan people from Guatemala (Pye, 1992). Thus, beliefs based on cultural practices can affect the way parents talk with their children. It is possible that cultural differences are also associated with children's language development. Callaghan, Rochat, and Corbit (2012), compared language acquisition in children living in Canada, Peru and India. Their results showed that even though comprehension vocabulary was not affected by cultural rearing practices, children's language production was faster in children living in Canada in comparison to children living in Peru and India.

*Reason 3)* SES may also be associated with differences in the time available for parent–child interaction and in the magnitude of other stresses on parents, which could shape parents' interactions with their children (Hoff, Laursen, & Tardif, 2002; Hoff & Tian, 2005; Vernon-Feagans et al., 2020).

In fact, in some cases, SES could also be considered an index of early adversity or even poverty that has long term effects on children's language outcomes. By the time they enter kindergarten, children from disadvantaged backgrounds differ substantially from their more advantaged peers in verbal and other cognitive abilities (Ramey & Ramey, 2004), and those early disparities that are predictive of later academic success (V. E. Lee & Burkam, 2002). These SES differences in language proficiency can still be seen in adults indicating that SES differences in language skills are robust and cumulative, and expand across the lifetime (Pakulak & Neville, 2010).

More work on the variables surrounding SES is needed to understand what factors are most related to language development. Moreover, it is possible that SES based on maternal education does not translate across all cultures and societies. A large study examined child development and growth in young children across socio-economic position in India, Indonesia, Peru, and Senegal (L. C. Fernald, Kariger, Hidrobo, & Gertler, 2012). In all countries, household wealth and maternal education contributed significantly and independently to the variance in the Extended Ages and Stages Questionnaire (EASQ) which was administered to parents of children aged between 3 and 23 months in the household, as well as to the variance in children length measurements (taken for all children between 0 and 23 months). This study shows that maternal education is still a relevant construct in other non-western cultures, including India.

Considered together, this work highlights children's different language experiences around the world. It also shows a gap in the literature documenting non-western children's language experiences as well as language skills and highlights the importance of measuring contextual variables that are often not accounted for in western populations. An environmental variable that has been shown to have an impact of children's development in western and non-western populations is early adversity. Children exposed to early adversities such as nutritional deficits or low parental education are at a high risk of delays in their cognitive development. An estimated 250 million children (about a 43%) in low and middle income countries fail to reach their developmental potential due to early adversity (Black et al., 2017).

Thus, it is important to better understand what contextual variables play role on language development and for that, we need to study different populations. We also need to better document what processes, language skills, and cognitive abilities influence the task of learning a new word across children around the world. The present project aims to fill this gap by measuring early language experiences and language processes in children from the UK and India.

## 1.3 The Present study

Better understanding of the relationships between children's early environment and their word learning abilities requires documenting early language experiences in different cultures and populations as well as measuring multiple language processes involved in word learning. This project was designed to do exactly that.

The aims of this project are: 1) Develop an early language processing task that includes several measures of language processing and test it with children living in the UK and India at different times in development (Chapter 2 and Chapter 4). 2) Measure the relationship between language home language input and language processing skills at different time points in both the UK and India (Chapter 3 and Chapter 5). These two aims are divided into four chapters that address the aims based on cultural context. We did not do direct cross-cultural comparisons, rather we looked at children in relation to their own population. The structure of the thesis is as follows.

Chapter 2 discusses the development of the Early Language Processing (ELP) task, an eye-tracking based measure of language processing, as well as evidence that this new task is able to measure different language processes in a sample of UK children. This study includes test-retest data from the same children at two time points when they were 15 to 27-months-old and when they were 28 to 36-months-old.

Chapter 3 describes how we investigated the relationships between early home language input gathered when children were infants (6 months of age) and toddlers (18 months of age), and the same children's language processing abilities on a subset of ELP test and retest data reported in Chapter 2.

In Chapter 4, we explain how the ELP task was adapted to be used in rural India to measure language processes in a totally different setting, language, and culture across development. This study includes test-retest data from the same children at two time points, when they were 15 to 27-months-old and when they were 28 to 36-months-old, mimicking Chapter 2. Here we also considered environmental variables such as SES based on maternal education.

Chapter 5 looks at the relationships between language input collected in Indian homes during infancy and toddlerhood, and the same children's language processing abilities measured with the ELP task reported in Chapter 5. SES based on maternal education was also considered in this sample.

Finally, Chapter 6 integrates the findings from the previous chapters and how they contribute to our understanding of early language development. Future directions and limitations of this work are discussed.

# Chapter 2

# Measuring Language Processes Early in Development with the new Early Language Processing Task

## 2.1 Introduction

Early language development is predictive of later language and cognitive abilities. For example, children's early language skills have been related to overall intellectual ability (e.g., Feldman et al., 2005) as well as the development of executive functions (e.g., Wade, Browne, Madigan, Plamondon, & Jenkins, 2014) and even academic success (e.g., Agostin & Bain, 1997). Speech perception, segmentation, and recognition skills measured in one-year-old children predict vocabulary measures in the second and third years of life (for a systematic review, see Cristia et al., 2014). Furthermore, children's expressive vocabulary and sentence complexity in preschool is predictive of literacy development (Scarborough, 2009). These findings suggest that early language abilities such as word processing speed, segmentation abilities, and recognition skills, together with comprehension and production vocabulary, support cognitive skill and intellectual functioning from early in development.

Interestingly, there is also a great deal of individual variability in children's early language ability. For example, when looking at children's vocabulary growth in Wordbank (an open database featuring data from parent-report vocabulary questionnaires from contributors around the world; Frank et al., 2017), we see that by 8 months, children are reported to understand between 2 and 56 words, and by 12 months between 21 and 185 words (i.e., deciles 0.10 and 0.90). Thus, even though some 8-months-old are reported to only understand 2 words, some of their peers understand even more words than infants who are 4 months older. There is in fact so much variability, that we can see that the ranges of words understood at 8 and 12 months largely overlap. This is also the case for word production (Frank et al., 2017). These early individual differences have also been associated with later language skill. Early delays in word learning predict group differences in vocabulary, syntax and verbal memory, in school age children (Rescorla, 2009).

Studying the mechanisms that support early word learning is crucial to better understand how they influence later language and cognitive skill. Moreover, measuring individual differences early in development can help to better understand how those differences emerge, and how they influence language development. However, while the literature shows that variation in early language development is a key predictor of later cognitive abilities, many of the studies in this literature use indirect measures of language development such as parental report of words understood and produced, or checklists of communication behaviours. Indirect measures are useful for tracking children's attainment of standard language milestones, and can be used to compare individual and group data on language development. However, these measures do not assess the underlying basic cognitive processes that support word learning and language development. Moreover, it is hard to assess language comprehension using parental report. What does it mean to *know* a word? How do you know your child knows the word "apple"? When parents are asked to say if their child understands "apple", they need to be able to think of situations in which their infants show understanding of those words and those situations can be ambiguous. This is why some studies have reported parents' underestimating or overestimating their children's abilities to understand words (especially in low-income families Roberts, Burchinal, & Durham, 1999; but see Reese & Read, 2000). Moreover, the concept of *knowing* a word might change with a child's age, as they are able to produce more words and their knowledge becomes overt.

A few studies in the literature have assessed early language processes using more direct measures, examining how individual differences relate to later abilities. A well established task to measure children's language processing abilities is the Looking While Listening (LWL) paradigm (A. Fernald et al., 2008). In this task, children see two images side by side (a target and a distractor) and hear sentences containing the target word such as "Look at the *target*". The measure of interest is children's efficiency in recognizing the target word, which can be assessed by either extracting "how much" the child looks at the target (usually proportion of looks), or "how fast" the child looks at the target (children's reaction time, RT). This paradigm has been used to measure several language processing abilities such as word comprehension and speed of processing; however, to our knowledge, direct measures of language ability have never been integrated in a single task to examine relations among measures.

The goal of this study is to create a task using the LWL paradigm that can be used early in development with individual children to directly measure multiple language processes. Such a task would provide better understanding of how language processes are related to one another and provide insight on children's language learning potential. To this end we present the Early Language Processing (ELP) Task. The ELP task is based on several well established tasks that have been shown to be predictive of later language skill: speed of language processing, online word comprehension, novelty biases, referent selection and retention of new words. We review these tasks and their predictiveness below. We then present test and retest data from the ELP, examining change over development in a sample of children living in the UK.

### 2.1.1 Speed of Word Processing

A well known measure using the LWL paradigm is speed of word processing (SoP), defined as how fast a child looks to a familiar image in response to a familiar spoken word when their first look was towards the distractor image (A. Fernald et al., 1998). A. Fernald et al. (1998) examined the time course of word recognition in infants ages 15 to 24 months, finding that efficiency of verbal processing increases dramatically over the 2nd year of life. Specifically, 15-month-old infants did not orient to the correct picture until after the target word was spoken, whereas 24-month-old started shifting their gaze to the correct picture before the end of the spoken word (see also A. Fernald, Swingley, & Pinto, 2001; A. Fernald et al., 2006; Zangl & Fernald, 2007).

Speed of word processing (also known as lexical processing or word recognition) has been associated with vocabulary size, such that children with faster word recognition have larger vocabulary sizes. In a longitudinal study of English-learning children, A. Fernald et al. (2006) explored the relationship between online speech processing efficiency and vocabulary growth during the second year of life. Speed of processing data was gathered at 15, 18, 21, and 25 months. The time course of eye movements in response to speech while looking to familiar images revealed significant increases in children's processing abilities over this period, reflecting better word comprehension. Moreover, both speed and accuracy of word recognition at 25 months were correlated with measures of lexical and grammatical development from 12 to 25 months; children who were faster and more accurate in online word recognition at 25 months, also showed a faster and more accelerated growth in expressive vocabulary across their second year. Thus, reaction time at 25 months was strongly related to lexical and grammatical development over the second year. In a follow up of this study, children who were originally tested as infants in their speed of processing abilities were assessed at 8 years on standardized tests of language, cognition, and working memory. Children's speed of spoken word recognition and vocabulary size at 25 months where both predictive of linguistic and cognitive skills at 8 years of age (Marchman & Fernald, 2008).

Individual differences in lexical processing predict not only long-term language outcomes, but short-term as well. A. Fernald and Marchman (2012) found that late-talking toddlers that had faster lexical processing abilities at 18 months were more likely to move into a normal range of vocabulary scores by 24 months, compared with late talkers that were less efficient in early speech processing. Similarly, the lexical processing speed of 18-month-old (adjusted age) preterm children was the strongest predictor of receptive vocabulary at 36 months of age, but uncorrelated with degree of prematurity or a composite of medical risk (Marchman, Adams, Loi, Fernald, & Feldman, 2016). In fact, speed of word processing at 18 months predicted receptive vocabulary, global language abilities, and non-verbal intelligence (IQ) at 4.5

years, even when controlling for socioeconomic status, gestational age, and medical complications of preterm birth. Importantly, speed of language comprehension remained uniquely predictive when also controlling for children's language skills at 18 months. Marchman et al. (2019) explored this relationship further by measuring both preterm and full-term children's vocabulary growth from 16 to 30 months, language processing speed at 18 months, and by accounting for a history of medical complications. Both preterm and full-term children displayed similar vocabulary trajectories up to 30 months of age, when birth group disparities began to emerge with preterm children showing slower language processing skills. Critically, language processing speed predicted expressive vocabulary size at 30 months. In preterm children, faster language processing speed predicted stronger outcomes regardless of number of medical complications, whereas slower processing speed and more medical complications predicted poorer outcomes. These results suggest that early differences, at least those observed in lexical processing efficiency, might have cascading consequences for language learning, which could be related to individual differences in language proficiency and even cognitive abilities observed in older children.

### 2.1.2 Online Word Comprehension

Clearly, there is good evidence of relationships between early lexical processing efficiency and later vocabulary size. It is also the case that vocabulary in infancy alone has been shown to be predictive of later language skills. In a study by Duff et al. (2015), pre-literacy vocabulary knowledge (i.e., between 16 and 24 months) assessed using parental report was predictive of later vocabulary, phonological awareness, reading accuracy and reading comprehension 5 years later (i.e., when children were between 4 and 9 years of age). Thus, it is possible that vocabulary in infancy is a platform for developing reading accuracy and reading comprehension skills. However, it is worth noting that the stability of vocabulary skills from infancy to later childhood in this study was too low to be sufficiently predictive of language outcomes at an individual level, and thus the conclusions should only be taken at a group level.

Studying early word comprehension is very challenging because, we aim to measure children's understanding of words that they do not yet say. Better understanding of infant's word comprehension is crucial because it provides the earliest window into children's understanding of word-referent relationships (Bates, 1993). Vocabulary checklists are powerful and well established tools that allow researchers to asses comprehension and production vocabulary sizes, however they do not tap into the cognitive mechanisms behind word comprehension abilities. In a language comprehension task there are a lot of processes involved such the strength of that word in memory and in the lexical network, general understanding of a category to recognise a word, etc. and thus, we need tools that measure these processes and how they contribute to children's overall ability to comprehend words.

In the last ten years, researchers have moved towards direct measures of children's lexical abilities by applying the LWL paradigm to the study of early word comprehension. This approach uses visual images to test children's knowledge of a word based on the looking patterns of the child, usually visual fixation or overall proportion of looks to target (e.g., A. Fernald et al., 2001). The downside of this work is the labor-intensive coding of the video data as they don't use automatised tools to measure child's gaze. This requires intensive coding resources and potentially a more limited sample size. In this context, the Computerised Comprehension Task (CCT) was created as a good alternative because it measures the child's performance using a direct and automatised task (Friend & Keplinger, 2003, 2008). The CCT is a touchscreen-based assessment that measures children's comprehension using children's touch as a response to a prompted word. Large image pairs appear on the screen and the child touches the target image in response to auditory prompts from an experimenter in which target word is embedded (e.g. 'Where is the shoe?' 'Touch the shoe.'). A significant contribution of this task is that it is administered in an engaging interface with easy data extraction, facilitating data collection in children up to 20 months (Friend & Keplinger, 2003).

The CCT builds on preferential-looking studies (e.g., A. Fernald et al., 2001) and picture book approaches (Ring & Fenson, 2000). It presents two pairs of images in a forced-choice format. The images represent different types of words (nouns, verbs and adjectives) that vary in frequency of occurrence in the typical receptive lexicon of infants (Dale & Fenson, 1996). Specifically, lexical targets were selected from the MacArthur-Bates Child Development Inventories (CDI: Words and Gestures and the CDI: Words and Sentences; Fenson et al., 2007). Based on those checklists, nouns, verbs and adjectives comprise about 75% (nouns 52.3%; verbs 18.8% and adjectives 5.7%) of infants' receptive vocabularies at 16 months of age (Fenson et al., 1994). In a study using the CCT task, directly assessed vocabulary comprehension in the 2nd year of life was also predictive of language skills during the 4th year of life, when children were in kindergarten (Friend, Smolak, Liu, Poulin-Dubois, & Zesiger, 2018). The authors found this pattern of results in English monolingual, French monolingual and French–English bilingual children. These results support the idea that early vocabulary may provide a foundation for later vocabulary and kindergarten readiness. A follow up study explored whether vocabulary comprehension measured using a direct task was as predictive as vocabulary measured using parental report (Friend, Smolak, Patrucco-Nanchen, Poulin-Dubois, & Zesiger, 2019). Results from this study showed that vocabulary comprehension measured with the CCT task was a stronger predictor of language skills than parent reported production measured with the MCDI in two linguistically and geographically distinct samples of American English and Swiss French children.

Some studies have started measuring both speed of word processing and online word comprehension in the same task using touch (e.g., Scaff, Fibla, & Cristia, in press; Smolak, Hendrickson, Zesiger, Poulin-Dubois, & Friend, 2021). Smolak et al. (2021) explored if decontextualized vocabulary (measured with the CCT task as the number of correct touch responses) and speed of word processing (measured as latency to fixate the target and latency to touch) at 2 years of age predicted vocabulary during the preschool period. Results reveal that at 2 years of age, vocabulary and visual response latency (but not haptic response latency) predicted vocabulary at 3 and 4 years of age. Further, decontextualized vocabulary remained a significant predictor when controlling for speed of processing, but not vice versa. This suggests interesting relationships between vocabulary, speed of processing and later language outcomes. For instance, the number of word–referent associations and the efficiency with which these are processed are important to vocabulary outcomes, but vocabulary seems to predict later skill more accurately in these age ranges. Relationships between speed of word processing and word comprehension had already been reported using visual paradigms. For example, A. Fernald et al. (2006) measured the relationships between online speech processing efficiency and vocabulary growth longitudinally. At 15, 18, 21, and 25 months children looked at pictures while listening to speech naming one of the pictures. Results of this study showed that the time course of eye movements in response to speech increased in the efficiency of comprehension over the 2nd year of life. Speed of word processing and accuracy of word recognition at 25 months were correlated with measures of lexical and grammatical development from the same children at 12 to 25 months. Moreover, children who were faster and more accurate in online comprehension at 25 months were those who showed more accelerated vocabulary growth in the 2nd year.

These studies provide evidence that both early vocabulary and speed of word comprehension are predictive of later language skills. Moreover, direct measures of vocabulary size seem to be more predictive of later vocabulary abilities. Finally, combining two predictive measures in a single task, such as speed of word processing and online word comprehension, allows researchers to examine how they influence each other. Particularly with previous research indicating that they might be associated. The present study builds on this literature to create an online task that uses children's looking patterns, rather than touch responses, to measure both speed of word processing and word comprehension early in vocabulary development. The created task adds measures of other early language processing skills shown to be critical in early vocabulary development. The advantage of administering these tasks using looking measures rather than touch measures is that this allows us to potentially test very young infants who might lack the skill to produce a touch after hearing a target word. Moreover, it allows tracking of looking patterns over time in trial. By examining how gaze changes in response to speech, we are able to study the time course of word recognition (e.g., A. Fernald et al., 2008; Mahr & Edwards, 2018).

### 2.1.3 Novelty, Referent Selection and Retention

To learn a word, children need to be able to find the referent, make a mapping, remember that mapping and integrate it with their previous vocabulary. We need to measure these processes to have a full understanding of young children's early processing abilities. In fact, children who rapidly recognize and interpret familiar words typically have accelerated lexical growth, providing indirect evidence that lexical processing efficiency is related to word learning ability. Lany (2018) found a relationship between speed of lexical processing and novel word learning in 18-month-olds and 30-month-olds. Children who were faster at recognizing familiar words were also more accurate at recognizing novel words when faced to a word learning task.

The task of learning a new word, however, is not that simple; there are several factors that matter. In word learning, children face referential ambiguity because when a new word is uttered the referent of that novel word must be selected from many possible objects present in the scene. Children are quite good at quickly mapping novel names to novel referents rather than to familiar ones (Mather, 2013). To do that, children rely on prior lexical knowledge and biases towards novelty. Children's use of prior lexical knowledge to determine the referent of a novel word is termed ”mutual exclusivity” (Markman & Wachtel, 1988). The idea is that, when children are presented with a familiar object and a novel one, if children know the label for the familiar one, they are able to map the novel label to the novel object by excluding the possibility that the novel name refers to the known object. Mutual exclusivity has been demonstrated in children from 17 to 30 months using multiple paradigms including 2-dimensional images on a screen (Bion et al., 2013) as well as 3-dimensional objects on a table (Horst & Samuelson, 2008). The use of mutual exclusivity to determine a referent has been shown to be driven by how well the child knows the familiar objects presented with the novel object. Children are able to disambiguate between a familiar object and a novel one when presented with a novel word if they have a strong association between the familiar object and the word that defines it. Studies manipulating children's knowledge of the objects, show the relevance of the strength of children's lexical representations because with weak familiar object knowledge children are not able to identify the target (e.g., Kucker & Samuelson, 2012). This indicates that children are able to map a new word into a novel object when that appears in the context of a highly familiar one. Thus, in mutual exclusivity, children bring their previous knowledge to bear in-the-moment to select the referent of a novel word, a dynamic process that contributes to building a lexicon (Kucker et al., 2018).

The ability to disambiguate is affected by children's knowledge of the familiar object, because highly familiar objects and larger vocabularies enhance children's mutual exclusivity abilities (e.g., Bion et al., 2013; Yurovsky et al., 2012). However, what makes a child attend to a novel object? Another explanation for children's disambiguation abilities during referent selection is children's attraction to novelty, a phenomena known as "novelty bias". The interaction between novelty driven attention and lexical knowledge has been evaluated in referent selection and retention tasks (RSR). Such tasks begin with a series of warm up trials during which children are asked to select each of three familiar objects by name (e.g., "Get the puppy"). On each experimental trial, children are presented with two of these familiar objects and one novel object. On novel name referent selection trials, children are asked for an object with a novel name (e.g., "Can you get the blicket?"). On familiar name trials, children are asked for a familiar object by name, now in the context of a novel object. Retention of novel word-object mappings is tested after a short delay by presenting children with three novel objects seen previously and asking them to get each, in turn, by name. Studies using the RSR task show the role of novelty bias as well as prior lexical knowledge in word learning, and how children's attention to novelty continuously changes over development (Kucker et al., 2018; Horst & Samuelson, 2008; Samuelson et al., 2017). Kucker et al. (2018), found negative associations between attention to novelty and retention of new word-referent links across individual 18-months-old children using 3-dimensional objects on the RSR task. In fact, at that age, novelty biases were so strong that children consistently selected the novel object with both known and novel names. This study also examined possible sources of bias though a computational approach, suggesting that when lexical knowledge is weak, attention to novelty supports in-the-moment behaviour but not learning (i.e., retention of the novel-object word association). In another study using a very similar version of the RSR task, 24 month-old children overcame the novelty bias and correctly selected a novel referent in response to a novel word, but they could not remember it after a 5 minute delay (Horst & Samuelson, 2008; Bion et al., 2013). By 30-months of age, however, children were able to overcome the novelty bias to select the correct referent and remember novel name–referent mappings over a time delay (Bion et al., 2013; Spiegel & Halberda, 2011).

These studies show the links between lexical knowledge and novelty biases and how those might relate to referent selection and retention abilities. It is important to measure novelty biases across development to better understand how novelty and familiarity impact word learning and the different language processing abilities that children use to learn new words. To do that, we need to develop tasks that incorporate those measures, so we can assess how they relate to one another. We have also reviewed literature showing relationships between speed of word processing and vocabulary abilities. Being able to combine all these measures in a single task would yield greater clarity into how different language processes contribute to the word learning task. Moreover, a direct measure of early language abilities though automatised procedures such as eye-tracking techniques, would enable researchers to use larger sample sizes facilitating a greater understanding of individual differences in early word learning.

### 2.1.4 The Present Study

The present study builds on prior work to create a new measure of early language abilities, the Early Language Processing (ELP) task, that integrates language processing measures of speed of word processing, online word comprehension, novelty biases, referent selection and retention of new words. The ELP task uses a remote eye-tracker and, thus, measures looking time and gaze trajectories as dependent measures. This allows testing of young participants using automatic, easy-to-implement protocols. Moreover, we designed the task to be portable, allowing the measurement of children's language skills in more naturalistic environments outside the laboratory. With an eye towards large scale employment in multiple populations, we wanted ELP to be not only portable, but efficient and adaptable as well.

We present data collected from a large number of children between 16 and 27 months of age using the ELP task, which we refer as the ”Test Group”. We examine the relation between our measures and prior similar tasks in the literature and look at developmental changes in our measures. We also collected data from a subset of the same children at a second time point between 30 and 32 months of age, which we refer as ”Retest Group”. These data allow examination of retention abilities in older children which should improve based on findings in the literature as well as an assessment of the reliability of the ELP measures across two time points.

ELP is an eye-tracking based looking task that lasts approximately 15 min long. It measures multiple processes that support word learning, combining four well-established measures: speed of word processing based on work from Fernald and colleagues (A. Fernald et al., 1998), word comprehension which gives a direct measure of a child's vocabulary size or word comprehension abilities based on the CCT task (Friend & Keplinger, 2003), referent selection or disambiguation which also includes a measure of novelty biases, and retention of new word-object mappings (Bion et al., 2013; Horst & Samuelson, 2008).

To overcome the limitations posed by young toddlers limited abilities, we used eye movements to capture children's looking patterns in response to audio-visual stimuli. This is a very simple response that allows us to tap the cognitive systems of interest at younger ages.

## 2.2 Methods

### 2.2.1 Participants

The final Test Group sample included 167 children aged 15- to 27-months-old (*M* = 20.23 months, *SD* = 3.03 months, 84 female; see age distribution on Figure A.1 in the Appendix). Participants had normal or corrected-to-normal vision. An additional 7 children were recruited but were not included in final analysis due to fussiness (2), technical problems (2) or not providing enough usable data (e.g., had noisy eye tracking data, 3). A subset of 76 children, the Retest Group, were tested again when they were between 28 and 36 months of age (*M* = 31.68 months, *SD* = 1.68 months, 35 female; see age distribution on Figure A.2 in the Appendix). Participants had normal or corrected-to-normal vision. Data from an additional 2 retest children are not included in the final analysis due to fussiness.

| **Sample Demographics; overall n = 161** | |
|---|---|
| **Age in Months** | |
| Mean (SD) | 20.43 (3.10) |
| Median [Min, Max] | 20.15 [15.00, 27.00] |
| **Ethnicity** | |
| African | 0 (0%) |
| Asian | 1 (0.6%) |
| Mixed | 9 (5.6%) |
| White | 147 (91.3%) |
| Not specified | 4 (2.5%) |
| **Parent 1 Education Status** | |
| Left School | 1 (0.6%) |
| GCSE/O levels equivalent | 11 (6.8%) |
| A levels or equivalent | 22 (13.6%) |
| Trade apprenticeship | 2 (1.2%) |
| Some university | 10 (6.2%) |
| Bachelor's Degree | 66 (40.9%) |
| Master's Degree | 29 (18.0%) |
| Doctorate or Professional Degree | 16 (9.9%) |
| Not specified | 4 (2.5%) |
| **Parent 2 Education Status** | |
| Left School | 1 (0.6%) |
| GCSE/O levels equivalent | 19 (11.8%) |
| A levels or equivalent | 29 (18.0%) |
| Trade apprenticeship | 14 (8.7%) |
| Some university | 9 (5.6%) |
| Bachelor's Degree | 48 (29.8%) |
| Master's Degree | 20 (12.4%) |
| Doctorate or Professional Degree | 11 (6.8%) |
| Not specified | 10 (6.2%) |

**Table 2.1:** Summary of sample demographics for ELP.

This project was reviewed and approved by the UK NHS Health Research Authority Ethics committee (Protocol ID: IRAS 196063; PI: John P. Spencer and ID: 211250 PI: Larissa K. Samuelson). Parents signed an informed consent form. Children received a small toy of their choosing and a t-shirt for participating. A subset of the data reported here are also part of a larger study examining the early precursors of executive function led by Prof. John Spencer.

Table 2.1 shows the sample demographics. We did not obtain demographic information (i.e., ethnicity, parental education and annual income) for 6 participants. The sample of children was 91.3% white, 0.6% asian, and 5.6% mixed race. 71.3% of mothers had completed a Bachelor's degree or higher. Mean family annual income ranged from £31,200 to £36,399.

### 2.2.2 Materials

The ELP task was presented on a 24-inch BenQ Zowie XL2430 (up to 144 Hz) monitor screen that was connected to a Gigabyte mini computer used to display the stimuli and a Lenovo laptop host that interfaces with the eye-tracker software running SR Research Experiment Builder, which we also used to program the task. Participants were seated on their caregivers lap or on a high chair, approximately 80 cm from screen. The eye to camera distance was about 50 cm - and the eyes were in line with the top part of the screen. The eye-tracker was positioned at the horizontal center of the screen. The eye tracker was an Eye-Link Portable Duo (SR Research, Ontario, Canada) in the remote setting. Both screen and eye tracker were placed on a table (together with the Gigabite mini computer and the Lenovo). Due to the portable aspect of this setup, we also allowed the experimenter and the laptop that monitored the experiment to be in the same room. We trained the experimenter to not distract, give feedback or engage with the participant during the task. The setup is shown in Figure 2.1.

A small target sticker was placed on participants' foreheads which allowed tracking of head (and eye) position even when participants moved or the pupil image was lost. The eye tracker was set to monocular recording such that it

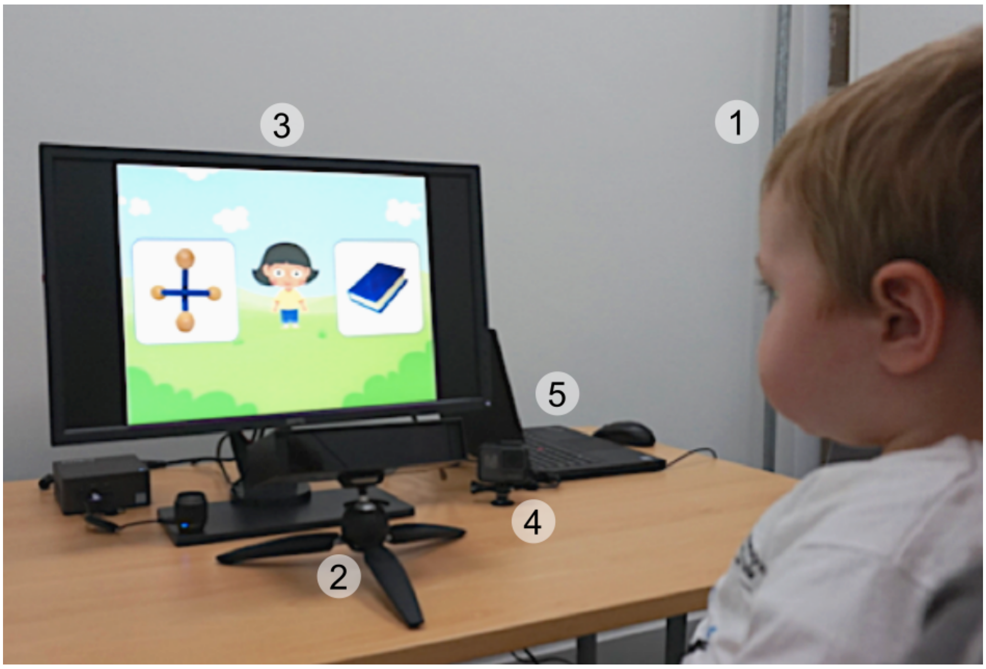


**Figure 2.1:** Portable ELP setup in the UK: 1) participant 2) eye tracker 3) screen 4) participant view camera 5) computer interfacing eye tracking software.

tracked the gaze position of a single eye using pupil and corneal reflections of an infrared light source. The sampling rate was 500 Hz. As part of the set-up there were two additional cameras in the room, one located on top of the monitor using a tripod, which recorded the participant's face (a GoPro model HERO5) and one located in the back of the room to record the experiment as it was presented on the monitor. These recordings were done to monitor and keep a record of the participant doing the task. Our portable setup also included a foldable silicon keyboard, a mini Xmi Pte Ltd portable speaker and a standard computer mouse.

### 2.2.3 Procedure

Before or after the ELP task as convenient, parents of participants completed an adaptation of the Oxford Communicative Development Inventory (OCDI, Hamilton, Plunkett, & Schafer, 2000) and also indicated which of an additional list of words their child understood and said. The additional words were those included on the ELP task that were not included in the standard OCDI. 2.2

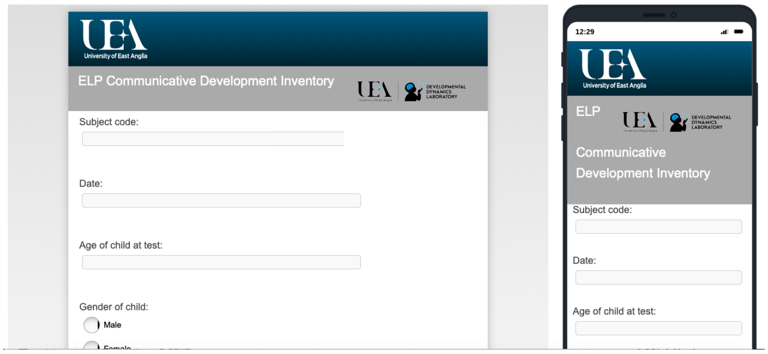


**Figure 2.2:** Visualisation of the OCDI online questionnaire adapted.

The ELP task began with a short clip of *Fantasia, 1995* (Disney). While this video played, the experimenter placed the small target sticker on the participant's forehead. Once the target sticker was in place, the tracking camera was adjusted so the distance from target to camera was approximately 50 cm. The experimenter adjusted the participants as many times as needed so they would be placed in the most optimal position and distance. After checking that the pupil and corneal reflection were visible on the camera, the calibration procedure began. During calibration, participants were shown a looming black and white geometric shape in five locations of the screen (middle, top, bottom, left, right) used to map raw eye position data to the camera image data and thereby allowing mapping of gaze position to the stimulus presentation. Following successful calibration, the experiment commenced. ELP was divided in two blocks separated by a 5 minute break. The calibration procedure was repeated after the 5 minute break.

### The ELP Task: Deisgn and Stimuli

Each ELP trial started by displaying two pictures on the screen for 2000 ms. Then, the screen was covered by a gray transparent filter and a gaze contingent cartoon appeared in the center. When the child looked at her, she named the target embedded in a carrier sentence such as ”Look, were is the (target)?”. At the onset of the target word, the gray filter disappeared and the child could clearly look at either the named image (target) or at the other one (distractor). The pictures remained for a 3200 ms response period and finally there was a reward which consisted of the cartoon character happily jumping up and down. This positive reward was always displayed (see the left panel of Figure 4.2 for the general structure of the ELP trial).

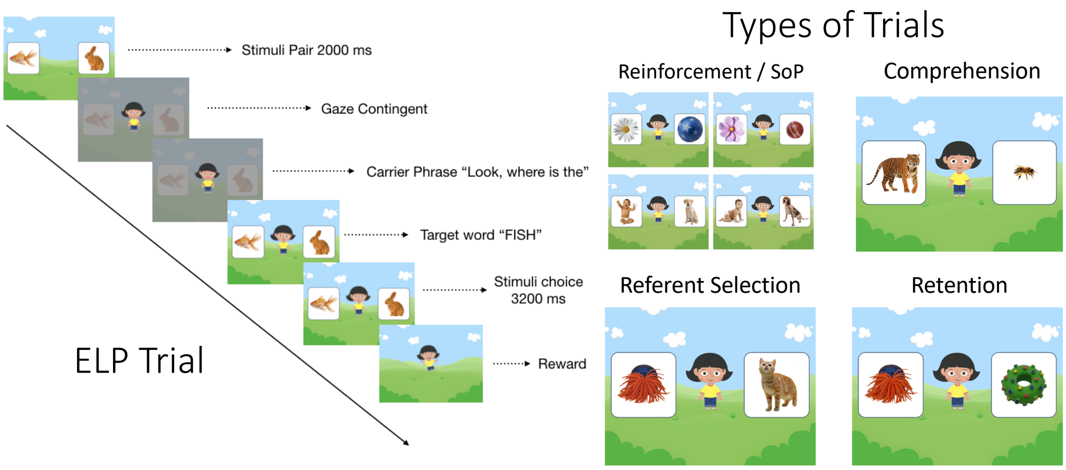


**Figure 2.3:** Trial schematic for the ELP task including trial types with examples.

The ELP task includes five measures of language processes: speed of word processing, comprehension vocabulary, novelty biases, referent selection and retention of new words. Those measures were incorporated in the task using four different types of trials (see examples of each trial on Figure 2.3):

- **Reinforcement or Speed of Processing Trials:** These trials contain two pairs of highly familiar nouns (*flower-ball* and *baby-dog*) that repeat 5 times each during the task. Two different sets of images were used on these trials to keep children interested in the task.

- **Comprehension Trials:** These trials include 41 pairs of nouns, verbs and adjectives varying in difficulty (see Table 2.2).

- **Referent Selection Trials:** These trials contain 8 word-image pairs with one well-known and one novel object. The novel was the target on 4 of those trials, and the familiar on the other 4. Half of the familiar images used on Referent Selection trials were extracted from Reinforcement trials (i.e., they were highly familiar nouns to which the child was exposed five times before seeing them in the context of a novel image). The other half were familiar nouns that also appeared on Comprehension trials (i.e., the child saw them only once before they appeared paired with a novel). This allowed us to manipulate image familiarity from highly familiar (images from Reinforcement trials) to less familiar (images from Comprehension trials).

- **Retention Trials:** These trials include image pairs of two previously-mapped novel objects. In the 4 retention trials, the child was asked to look to the novel word-image they saw during Referent Selection trials. Children were also exposed to a distractor novel image which appeared as a foil on Referent Selection trials where the target was the familiar. Thus, both images on retention trials had been seen previously.

Familiar words in the task were selected based on frequency of appearance in several British English natural language corpora from children between 0.4 and 5 years of age (Manchester Corpus Lieven, Salomo, & Tomasello, 2009; Wells Corpus Wells & Bridges, 1981; Quigley-McNally corpus Quigley & McNally, 2013 including only typically developing children; total number of

| Pairs | Easy | Moderate | Difficult | Total |
|---|---|---|---|---|
| **Nouns** | 7 | 7 | 8 | *22* |
| **Verbs** | 4 | 4 | 3 | *11* |
| **Adjectives** | 3 | 3 | 2 | *8* |
| **Total** | *14* | *14* | *13* | *41* |

**Table 2.2:** Word pair distribution on ELP by type and difficulty

types = 9710; total number of tokens = 766535) from the CHILDES database (MacWhinney, 2000). All the corpora were combined into a single one and used to calculate word types' frequencies. Frequency of appearance in children's language input was used to classify words into easy, moderate and difficult. For example, we selected the easy nouns for our task from a list of nouns that appeared between 1800 - 200 times in the corpora (see Table 2.3 for frequency ranges used in each word type and difficulty). This method was based on previous studies using the same approach (Maniel, 2016; Fibla, Maniel, & Cristia, 2016; Scaff et al., in press).

| | Easy | Moderate | Difficult |
|---|---|---|---|
| **Nouns** | 1800 - 200 | 200 - 100 | 100 - 3 |
| **Verbs** | 1000 - 100 | 100 - 30 | 30 - 4 |
| **Adjectives** | 1100 - 200 | 200 - 25 | 25 - 3 |

**Table 2.3:** Difficulty classification based on frequencies of word type

The audio stimuli were recorded using a female native speaker of British English. We asked her to speak in a child directed manner. Audio recordings were recorded using the GarageBand application by Apple with a mac OS, which includes a function to remove background noise. Stimuli were later extracted from the recordings and processed using Praat scripts (Boersma, 2001). We added silences at the edges of each sound file (0.01 s on each edge), and we normalised the sound intensity (i.e., amplitude). We recorded several examples of each word with its carrier sentence. Per each recording we extracted the total duration, root-mean-square pressure (i.e., the square root of the average of the square of the pressure of the sound signal over a given duration), the intensity in decibels and the average, minimum and maximum pitch. Those measures helped inform our selection of the best and clearest examples of each of the words we recorded.

Familiar images in ELP were selected from several open sources and matched in salience, colour and complexity. We asked several adults to match each word with one or more image candidates. Images that did not accurately match the word were replaced. As we will show in Chapter 4, the ELP was adapted to another language and culture. During the adaptation process some of the original words and images that did not suit the other culture were replaced with the goal of developing a task that would be comparable across sites.

Novel words and images were selected from the NOUN database (Horst & Hout, 2016). Two of the novel words selected were mono-syllabic *whilp, bink* and the other two bi-syllabic *koba, teebu*. None of the four words started with the same consonant and all contained different vowels. The selected words were not currently in use in other studies in the laboratory. Novel images were matched in salience and colour.

All images were scaled and processed with the GIMP software (The GIMP Development Team, n.d.). Animations such as the character moving its mouth were done using PowerPoint and exported into a video format. The ELP task was programmed using Experiment Builder (SR-Research, Ontario, Canada).

The ELP task consisted of two blocks separated by a 5-minute break (retention interval) during which children could either stand up or watch a short movie on the screen (Piper, a 2016 computer-animated short film produced by Pixar Animation Studios). Children started ELP with a first block that contained 5 reinforcement trials mixed into 20 comprehension trials followed by 8 referent selection trials (4 in which the target was the novel, and 4 in which it was the familiar). The different types of trials in block 1 were presented in the following order: 2 reinforcement trials, 7 comprehension trials containing easy and moderate nouns verbs and adjectives, 1 reinforcement trial, 8 comprehension trials with easy and moderate nouns, verbs, and adjectives, 1 reinforcement trial, 5 comprehension trials with easy and moderate nouns, verbs, and adjectives, 1 reinforcement trial, and 8 referent selection trials. After the 5-minute break and the second calibration (in case the child stood up during the break or moved from the initial position), the second block started. Children then were exposed to 5 reinforcement trials mixed with 4 retention trials and 20 comprehension trials. Trials in block 2 were presented in the following order: 2 reinforcement trials, 4 retention trials, 1 reinforcement trial, 5 comprehension trials containing moderate and difficult nouns, verbs, and adjectives, 1 reinforcement trial, 8 comprehension trials with moderate and difficult nouns, verbs, and adjectives, 1 reinforcement trial, and 7 comprehension trials with moderate and difficult nouns, verbs, and adjectives (see Table 2.4 for a summary of the ELP trial structure). For Comprehension trials, the first block only contained easy and moderate words, whereas the second block only contained moderate and difficult words. This meant that the task increased in difficulty as the child went through it. This helped ensure that even younger children would complete most trials in the first block.

In each block, word order was pseudo-randomised to ensure that the target did not appear on the same side of the screen more than two trials in a row, and that the word type/difficulty did not repeat more than two trials in a row. Referent selection and Retention trials were randomised separately but followed the same criteria such that the same word type would not appear more than two times as the target (for Referent Selection trials), and that the target would not appear on the same side more than two times (for both Referent Selection and Retention trials). Thus, we had two fixed pseudo-randomised ELP versions (order 1, order 2). To keep the task short, for each image pair, children were only asked for one of the images (but not the other). Thus, we created two different target word versions (A, B). For example, in the word pair *cat* - *fish*, order A asked the child to look at *cat*, and order B asked the child to look at *fish* (but in order A *fish* was never the target and in order B *cat* was never the target). This meant that the ELP task had four different versions based on target word and randomisation: A1, A2, B1 and B2. We tested approximately the same number of children in each order and checked for possible order effects in our analyses.

| **ELP Structure** | |
|---|---|
| **Block 1: Reinforcement/SoP, Comprehension and Referent Selection** | |
| Reinforcement/SoP<br>(highly familiar nouns) | 5 trials (mixed) |
| Comprehension<br>(easy and moderate nouns, verbs and adjectives) | 21 trials |
| Referent Selection<br>(pair of familiar and novel word-image) | 8 trials |
| *5 min break with animated video* | |
| **Block 2: Reinforcement/SoP, Retention and Comprehension** | |
| Reinforcement/SoP<br>(highly familiar nouns) | 5 trials (mixed) |
| Retention<br>(two novel images, one previously paired with a novel word) | 4 trials |
| Comprehension<br>(moderate and difficult nouns, verbs and adjectives) | 20 trials |

**Table 2.4:** ELP task structure. For Reinforcement/SoP trials, (mixed) indicates that those trials where mixed with the other trials of the same block.

The ELP aesthetics were designed in a child friendly way: the background was a picture of a field and the attention getter was an animated character programmed in a way that mimics talking to the child after the child look at it. It also provided positive feedback at the end of each trial by jumping up and down (background and character design are based on the Ipad App from Alejandrina Cristia, see `https://github.com/alecristia/mandy_newplugin`). This design has been previously used in a portable tablet-based vocabulary test in France (Fibla et al., 2016; Scaff et al., in press) and Argentina (Rosemberg & Alam, 2021), to measure comprehension vocabulary in toddlers from different socio-economic backgrounds.

The task used for the 30-month retest remained largely the same, with the exception that a new set of novel words and novel objects were selected from the NOUN database (Horst & Hout, 2016). This was to ensure the novel objects were novel, and not remembered from the prior test session. Again, two of the new novel words were mono-syllabic *foope*, *bem* and the other two bi-syllabic *tannin*, *osip*. The four words did not start with the same consonant and they all contained different vowels. They were also not in use in other studies in the laboratory. Novel images were matched in salience and colour.

All visual and audio stimuli can be accessed on the Project OSF site `https://osf.io/yczgj/`.

### 2.2.4 Analysis Method

The eye-tracking data were pre-processed using Data Viewer (SR-Research, Ontario, Canada). Trials were segmented into periods of interest (IP) using message-based events. Areas of interest (AOI) were set to be 50% bigger than target objects to account for calibration errors and drifts in the eye tracker. Sample reports were exported and raw gaze position was further analyzed using the statistical package R (R Core Team, 2017), as well as eyetrackingR

(Dink & Ferguson, 2016), an R package designed to work with eye-tracking data. A common measure in eyetracking studies of word recognition is an accuracy growth curve (also called Growth curve analysis – GCA; Mirman, 2014). The growth curve measures how the probability of fixating on the target changes over time. We computed this growth curve using eyetrackingR. Looking to the target side and the distractor at each point in time during the trial was aggregated into 100ms time bins allowing calculation of the proportion of looks to the target. We ignored offscreen looks or looks out of the AOIs when computing this proportion. Only trials with more the 60% of looking data were included in the analysis.

Out of the total ELP trial, our analyses focus on two windows of interest: one during the familiarisation phase, and the other during the test phase. We chose these two windows of interest with two objectives. First, we wanted to measure if children had a preference for any of the images before hearing the target word, that is, during familiarisation. Second, we wanted to measure if children looked at the target image after hearing the target word during the test phase. Looking data during familiarisation included looks towards the two images before hearing the target. To allow for the best possible statistical modelling of these time series data, the looking data from the first 300ms of the familiarisation phase of the trial was trimmed to reduce noise. Looking data from the test phase, focused on a window of interest that went from word onset to 1800ms after onset, consistent with previous studies. This criteria is based on previous literature suggesting that 24-months-old children shift their gaze to the correct picture before the end of the spoken word, in contrast to 15-months-old that do not orient to the correct picture until after the end of the target word (e.g., A. Fernald et al., 2008, 2001). Since we had a large age range, we wanted to take into account looks to target from word onset rather than word offset, since that would capture age effects in processing abilities. Looks across the ELP trial, as well as the different parts, are shown in Figure 2.4.

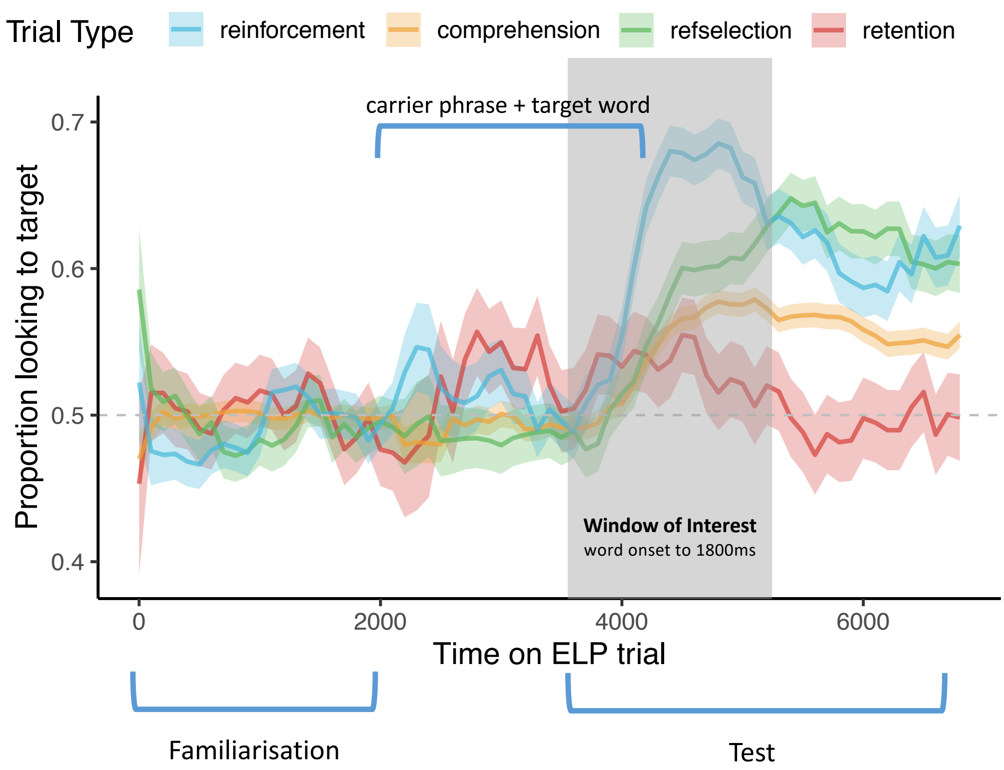


**Figure 2.4:** Looking patterns on the ELP trial plotted using proportion looking to target by time in the ELP trial. Colours indicate the different ELP trial types. The ELP trial parts are defined and, highlighted in grey, the window of interest from word onset to 1800ms after offset.

We had four specific questions we wanted to address: Q1) as a group, how did children perform on the different ELP measures on ELP Test (approximately 18 months); Q2) how did individual children perform across ELP measures at ELP Test (approximately 18 months); Q3) for those children who did ELP at both time points (Test and Retest), how did ELP performance at 18 months relate to performance at 30 months at the group level; Q4) and how did ELP performance at Test relate to performance at Retest at the individual level.

To answer question Q1, we run a first initial model that included all trial types. After, we focused on each trial type in particular. Those follow up models mimicked the structure of the initial model which was build as follows: proportion of looks to the target through time were fit with a binomial hierarchical model estimated with a Laplace approximation using the glmmTMB package (Brooks Mollie et al., 2017) and eyetrackingR (Dink & Ferguson, 2016) in the statistical package R (R Core Team, 2017). The model was fit with quartic orthogonal polynomials of the time term following the growth curve analysis approach (GCA) (Mirman, 2014), that is, the data were modelled with Time, time squared, up to time to the power 4, but scaled and centred so as to not be correlated with one another. In addition, the model contained fixed effects of Age in months represented as a continuous variable and Trial Type which included all ELP types of trials. Each of the time terms were nested as a random effect within participant, along with allowing each participant a random intercept for a maximally-specified model. The model was fit with Age, Gender and Trial Type. The model was then simplified using the Akaike information criterion (AIC), an estimator of prediction error and, therefore, relative quality of the statistical models, using the Anova function of the R package (Wagenmakers & Farrell, 2004). Because Gender did not show any consistent results nor improved AIC values (i.e., Anova comparisons between a model that included gender and a model that did not were not significant), Gender was removed from the models. Models were also tested using the DHARMa R package (Hartig, 2021), which creates readily interpretable scaled (quantile) residuals for linear mixed models, as well as plot and test functions for typical model miss-specification problems (e.g., over/underdispersion, zero-inflation, and residual spatial and temporal autocorrelation).

Some follow up models looking at particular ELP trials had a more complex random structure in which an additional variable (e.g., word) interacted with participant. Thus, in this case the random structure contained each of the time terms nested as a random effect within participant, along with allowing each participant a random intercept plus random variation in intercept among participants within that extra variable. We report particular details in the pertinent Results section.

The same modelling approach was used to answer Q3 but, because in this case we modelled ELP data at both time points (Test and Retest). The model also included Test Type (Test at 18 months versus Retest at 30 months) as fixed effect as well as a part of the random effects structure (interacting with participant). Thus, in this case the random structure contained each of the time terms nested as a random effect within participant, along with allowing each participant a random intercept plus random variation in intercept among participants within test type. This initial model was used to assess overall differences between ELP measures (Test versus Retest) in a single model. The same model (with the same fixed effects and random effects structure) was used to look at proportion of looks to the target through time, split by each ELP trial (or measure) at both Test (approx. 18 months) and Retest (approx. 30 months), as well as to control for image preference during familiarisation across the different ELP trials.

Variations from these two initial models are detailed in the pertinent section in Results. Any variation from the standard approach has been included because it substantially improved the model fit. Changes from the original model were assessed using the AIC criterion and Anova comparisons (see Wagenmakers & Farrell, 2004), as well as DHARMa plot and test functions (Hartig, 2021), for typical model miss-specification problems using R package. These were the same methods applied to evaluate the best initial model.

For each model, the effect of each parameter was assessed with an F test, in particular, we used the ANOVA function from the car R package (R Core Team, 2017), which tests whether the model terms are significant. All the reported effects and interactions are those that remained after using this method.

To answer Q2 and Q4, we used correlation analyses. We computed the overall mean proportion of looks to target on the test phase of ELP, in particular during the window of interest (from word onset to 1800 ms). We used eyetrackingR (Dink & Ferguson, 2016) to compute this proportion. Q2 focused on individual performance across ELP measures and, thus, we ran a set of correlations between the different ELP trial types at ELP Test at 18 months. We also considered OCDI. To correct for multiple comparisons, we set a more conservative criteria and only considered effects with a significance level smaller than 0.01 (sig.level <0.01). Q4 looked at individual performance across both ELP observations and, thus, we ran a set of correlation analyses to measure relationships between ELP Test at 18 months versus ELP Retest at 30 months. Correlations were run in the R package (R Core Team, 2017).

## 2.3 Results

Our participants did ELP both at approximately 18 and 30 months of age (test-retest). In this section, we examine the relations between the different ELP measures at both time points. We first present results on the ELP data collected at ELP Test (18 months), followed by results on ELP data at both Test and Retest (18 and 30 months comparisons). We use both GCA approach and correlation analyses to measure performance at the group level and at the individual level on ELP Test at 18 months, and performance at the group level and at the individual level across both time points (18 and 30 months).

### 2.3.1 ELP Test at 18 months

Here we report results in answer to Q1: how did children perform as a group in the different ELP measures when they were approximately 18 months? We then examine Q2: how did individual children perform across ELP measures at 18 months?

In a first big model, looking proportions were modelled following the GCA approach with a hierarchical binomial model to examine the effects of Trial Type, and Age (in months) over Time in the test phase of the task. The model utilized orthogonal cubic polynomials of the Time term to capture the model fit (Mirman, 2014), and included Time and participants in the random structure as described in the analysis method. The aim was to assess if the task is sensitive to the different language measures.

Results show main effects of the linear and quadratic Time terms, Age and Trial Type. There was also a significant 2-way interaction between Age and Trial Type and an interaction between Age and the linear and quadratic Time terms. Next, there was a 3-way interaction between the linear and quadratic Time terms, age and Trial Type (see Table 2.5 with F values at the end of this chapter). These results indicate that children's looking patterns change over Time in the ELP task as a function of Age and ELP measure.

The model fit to the raw data can be seen in Figure 2.5. As can be seen in the figure, there was an increase in the proportion of looks to the target on Reinforcement trials over age (see blue lines). This improvement over age was also evident on Referent Selection (green lines) and Comprehension trials (yellow lines). Performance on Retention trials (red lines) was variable at the younger ages, but looked to be reliably above chance for the oldest age group.

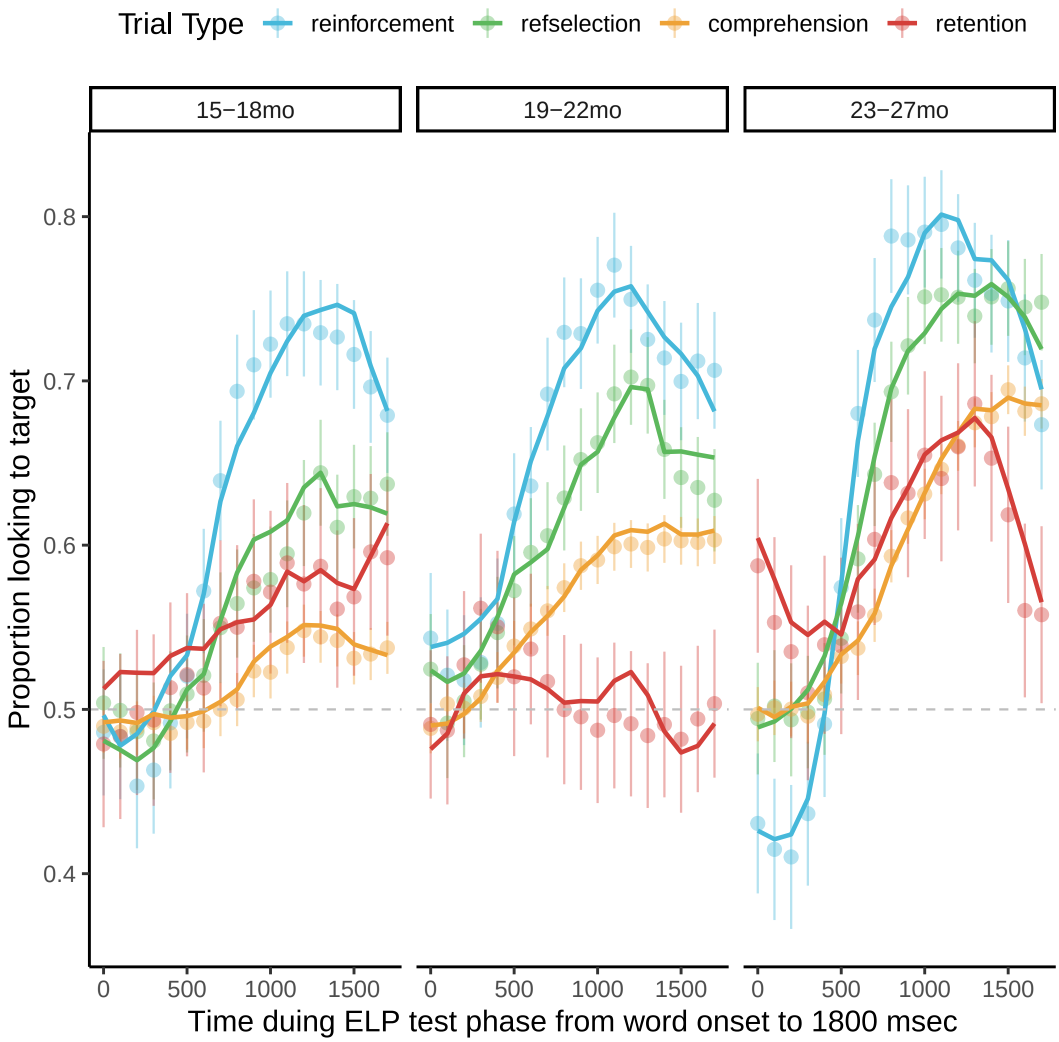


**Figure 2.5:** Model predicted proportion looking to target by Trial Type by age. Grey dashed line depicts chance performance (0.50). Age in months was split in three age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

To further examine the relationship between our ELP measures and independent measures of vocabulary development, we ran the same model using OCDI scores (comprehension or production) as a predictor variable instead of Age in months. The three models using age, OCDI comprehension and OCDI production (sub-sample of 116 18-month-old children) showed very similar AIC scores (Model with Age *AIC* = 151574.4 see F values on Table A.1; Model with OCDI Comprehension *AIC* = 150660.4 see F values on Table A.2; Model with OCDI Production $AIC$ = 150942.8 see F values on Table A.3), with the model using comprehension as a fixed effect showing the best fit to the data with smaller AIC values. The three measures (Age, OCDI Comprehension, OCDI Production) are correlated with each other (Age and OCDI Comprehension t = 217.13, r = 0.654, p<.000; Age and OCDI Production t = 290.05, r = 0.756, p<.000; OCDI Comprehension and OCDI Production t = 315.47, r = 0.782, p<.000). Reflecting this, the differences between the models are extremely small. Moreover, OCDI scores could potentially be biased due to being a parental report measure. Thus, we used Age treated as a continuous variable in the following analyses (see Appendix for Figure A.3 using OCDI Comprehension instead of Age).

Since the overall model showed evidence of differences across trials types, we modeled each trial type separately in more detail.

### ELP Reinforcement

Looking proportions were modelled using a very similar hierarchical binomial model as in the overall model but with an additional fixed effect Repetition Pair Count that modelled the repetition number of the reinforcement pair – a count of how many times the pair of highly familiar images had appeared in the task. We separately modeled both the test phase and the familiarization phase of the reinforcement trials. This was to assure the children did not have a systematic preference for any of the presented images, and that the possible effects found during the test phase reflect the language processes that occur after hearing the target word.

Results of the familiarisation phase model showed a main effect of the quadratic Time term and a interaction between the quadratic Time term and Age (see Table 2.6 for F Values). As can be seen in Figure 2.6, looking to the two images during the familiarization phase was roughly equal, oscillating around chance, with some small age-based differences. Thus, there are no clear biases in looking during familiarisation that would explain differences at test.

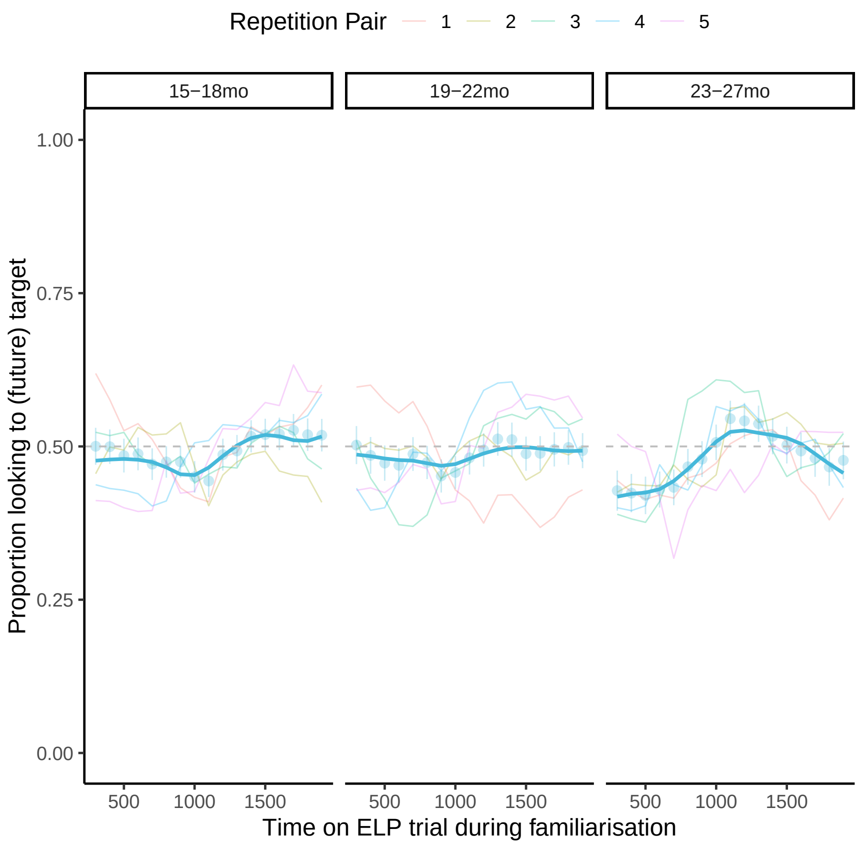


**Figure 2.6:** Model predicted proportion looking to target on Reinforcement trials by Age and Repetition Pair Count during familiarisation. Grey dashed line depicts chance performance (0.50). Age in months was split in three age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

To assess performance after the target word was spoken, the same hierarchical binomial model was applied to looking data from the test phase of the Reinforcement trials. Results show a main effect of the quadratic Time term, a 2-way interaction between the linear and quadratic Time terms and Age, a 2-way interaction between the quadratic Time term and Repetition Pair Count and a 3-way interaction between the quadratic Time term, age and Repetition Pair Count (see F values on Table 2.7). As can be seen in Figure 2.7, there was a clear effect of Age and of Repetition Pair Count: children's looking patterns change over the course of the trial as a function of their age, and the number of times they had been exposed to a particular word pair. Younger children tend to look at the target and stay there for longer than older children, who quickly look to the target but release fixation earlier. This suggests that older children know reinforcement words better than younger children as they are highly familiar. Regarding repetitions, children tend to look more reliably to the target as the number of repetitions increases.

To relate reinforcement trials to other ELP measures we extracted the following measures: a) overall proportion looking to target in the window of interest, b) model coefficients that express the rate of change looking to target in the window of interest (the linear time term), c) reaction time of first look to target. We explored these three indexes because they measure different aspects of word processing on the reinforcement trials. Overall proportion shows accuracy of children's looking. The model coefficient based on the liner term indicates the steepness of the growth curve – how fast and how much children look at the target over the trial. RT of first look shows how fast children looked at the target after being prompted with a word.

We did not find relationships between the three Reinforcement indexes. Neither between the Reinforcement measures and OCDI scores (see Appendix for a correlation plot between those measures including *r* values Figure A.4 and Figure A.5).

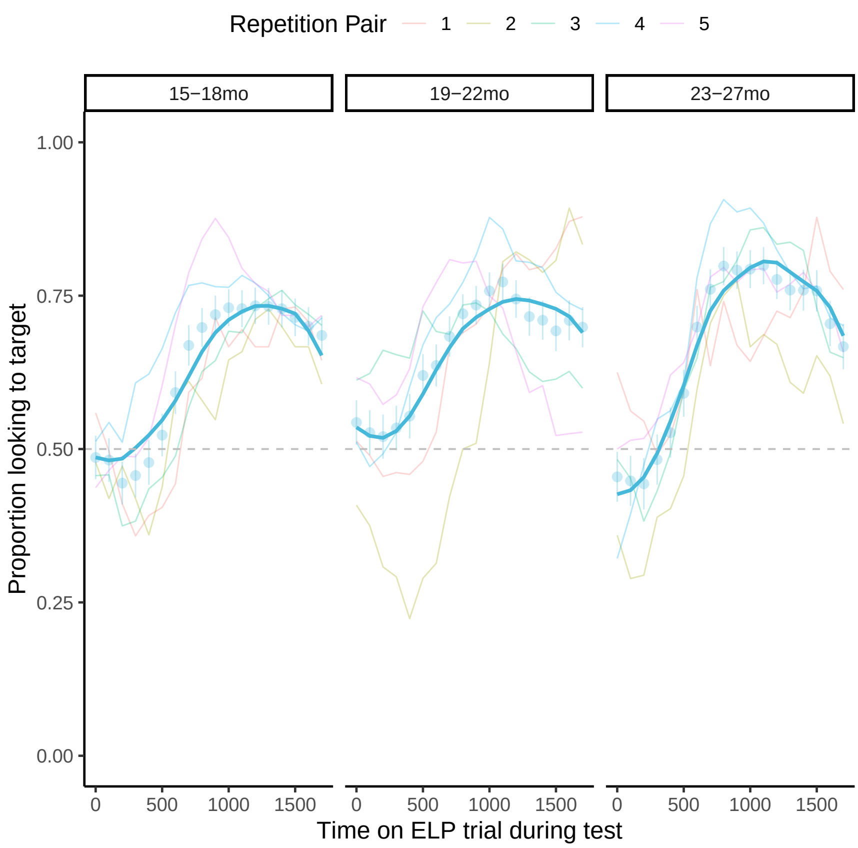


**Figure 2.7:** Model predicted proportion looking to target on Reinforcement trials by Age by Repetition Pair Count during test. Grey dashed line depicts chance performance (0.50). Age in months was split in three age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

### ELP Comprehension

Looking proportions were modelled following the same GCA approach with a hierarchical binomial model to examine effects over time during familiarisation and test phases on comprehension trials, as well as effects of age (in months).

A first binomial model included all word types and difficulties collapsed, to see overall comprehension effects. Thus, looking proportions were fit with Time up to the fourth term and Age. The model included Word, Time and participant in the random effects structure (with word and participant interacting with each other). Results from this model from the familiarisation phase showed no effects (see F values on Table 2.8). As can be seen in Figure 2.8, children show chance levels of looking, indicating no systematic preference for a particular image or side.

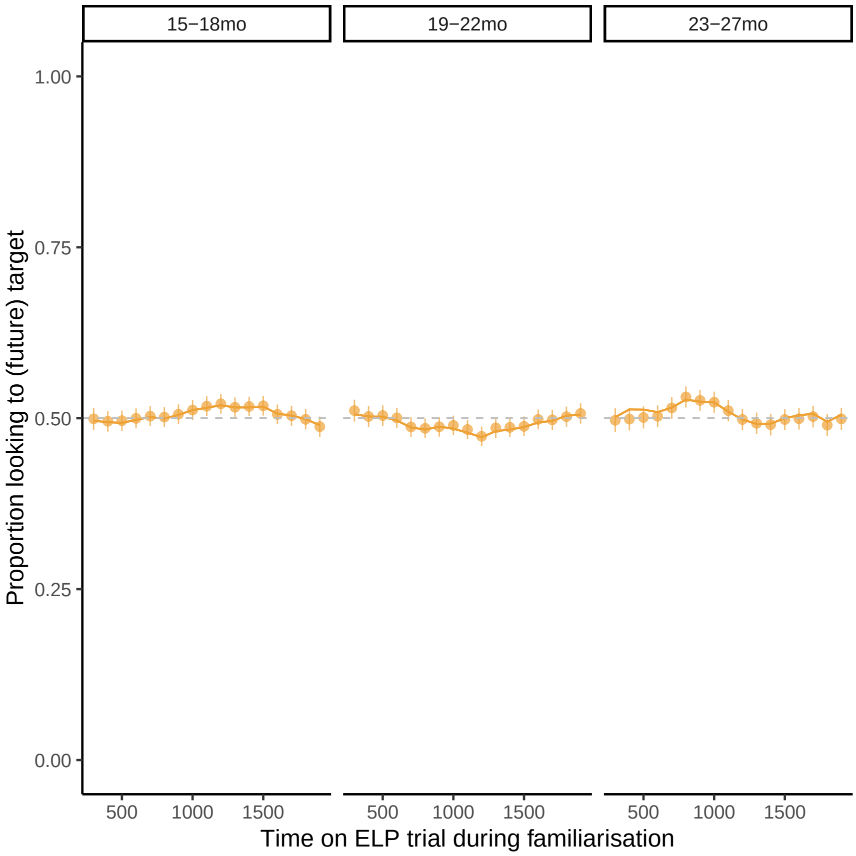


**Figure 2.8:** Model predicted proportion looking to target in Comprehension trials overall by Time by Age during familiarisation. Grey dashed line depicts chance performance (0.50). Age in months was split in three age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

Results from the test phase of Comprehension trials using the same binomial model as in the familiarisation show main effects of the linear Time term and Age in months, and a 2-way interaction between the linear Time term and Age in months (see Table 2.9 with F values). This indicates that children's looks toward the target increase over time in trial as a function of age. As can be seen in Figure 2.9, older children look more, earlier and faster to the target than younger children. However, this model looks at all comprehension trials collapsed across types and difficulties. Thus, we conducted additional analyses to examine, for instance, whether young children find some words hard, resulting in less looking towards the target. In particular, the next models looked at different comprehension trials to explore the effects of word type and difficulty. Because ELP has more nouns than verbs and adjectives, we first only looked at nouns split by word difficulty (easy, moderate and difficult) and the at adjectives versus verbs collapsed across word difficulties.

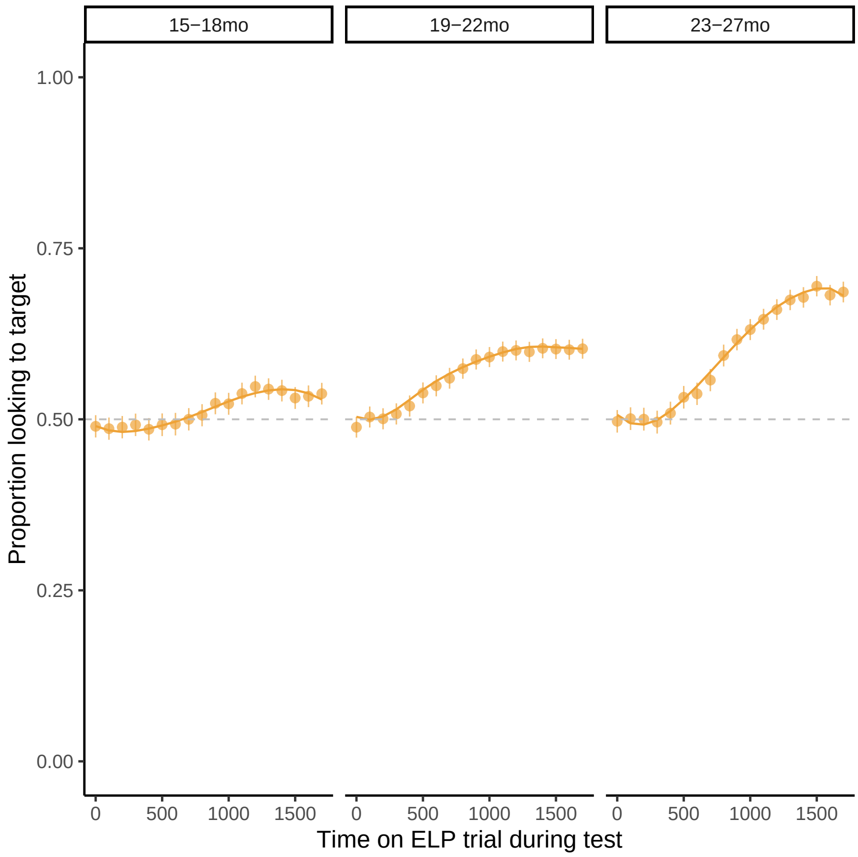


**Figure 2.9:** Model predicted proportion looking to target in comprehension trials overall by time by age during test. Grey dashed line depicts chance performance (0.50). Age in months was split in three age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

The noun data was modeled using the GCA approach in a model with only nouns with fixed effects of Word Difficulty, Age, and Time up to the third term. Word Difficulty interacted with Participant and Time in the random effects structure. This model showed a main effect of Word Difficulty, a 2-way interaction between the linear Time term and Word Difficulty, a 2-way interaction between Age and Word Difficulty, and a 3-way interaction between the linear, the quadratic and the cubic Time terms, Age and Word Difficulty (see Table 2.10 with F values).

The looking patterns can be seen in Figure 2.10, reflecting children's word knowledge as a function of word difficulty and age from 15 to 27 months. These results show that, overall, children look differently at nouns with different difficulty levels as a function of age. As expected, the rate of looking towards nouns over the course of the trial changes as a function of noun difficulty with more looks towards easy and moderate nouns in comparison to difficult nouns. Also, the amount and speed of looking to the target is moderated by child's age. Older children look faster and longer to the target. Moreover, with age, children get better at quickly recognising the target when this is an easy or moderate noun, but they have more trouble recognising difficult nouns.

Given the limited number of verbs and adjectives we tested, we only compared verbs and adjectives collapsing across word difficulties. Following the GCA approach, a third binomial model was fitted to the subset of the data containing looks towards verbs and adjectives on comprehension trials. The model structure included Word Type (verb versus adjective) by Age by Time as fixed effects. The random effects structure included Time and Participant interacting with Word Type. The model results (see F Table 2.11) showed a main effect of the linear Time term, a 2-way interaction between the linear

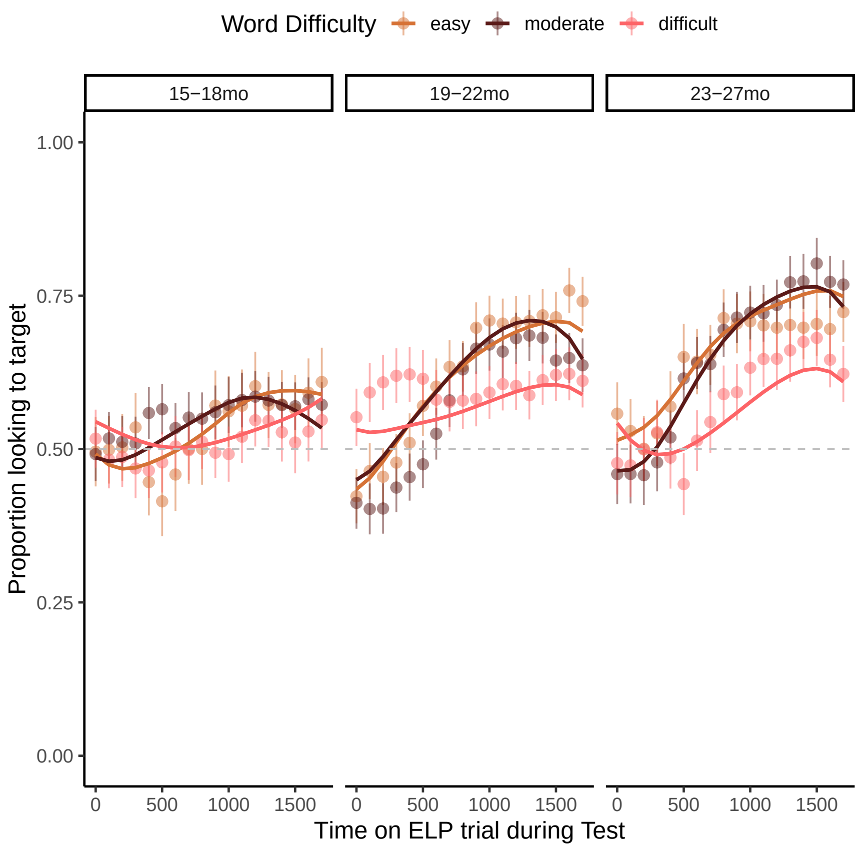


**Figure 2.10:** Model predicted proportion looking to target in Nouns from Comprehension trials split by Word Difficulty by Time by Age during test. Grey dashed line depicts chance performance (0.50). Age in months was split in three age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

Time term and Age, a 2-way interaction between the linear Time term and Word Type, a 2-way interaction between Age and Word Type, and 3-way interactions between the linear and quadratic Time terms, Age, and Word Type.

There was a strong age effect with older children looking more and faster towards the target for both verbs and adjectives. As can be seen in Figure 2.11, this effect is more pronounced for adjectives, because children look more and faster to adjectives than to verbs (in fact, older children seem to be able to recognise adjectives as well as nouns; see comparison data shown in red). Only older children show looks greater than chance levels towards verbs, mostly at the end of the window of interest.

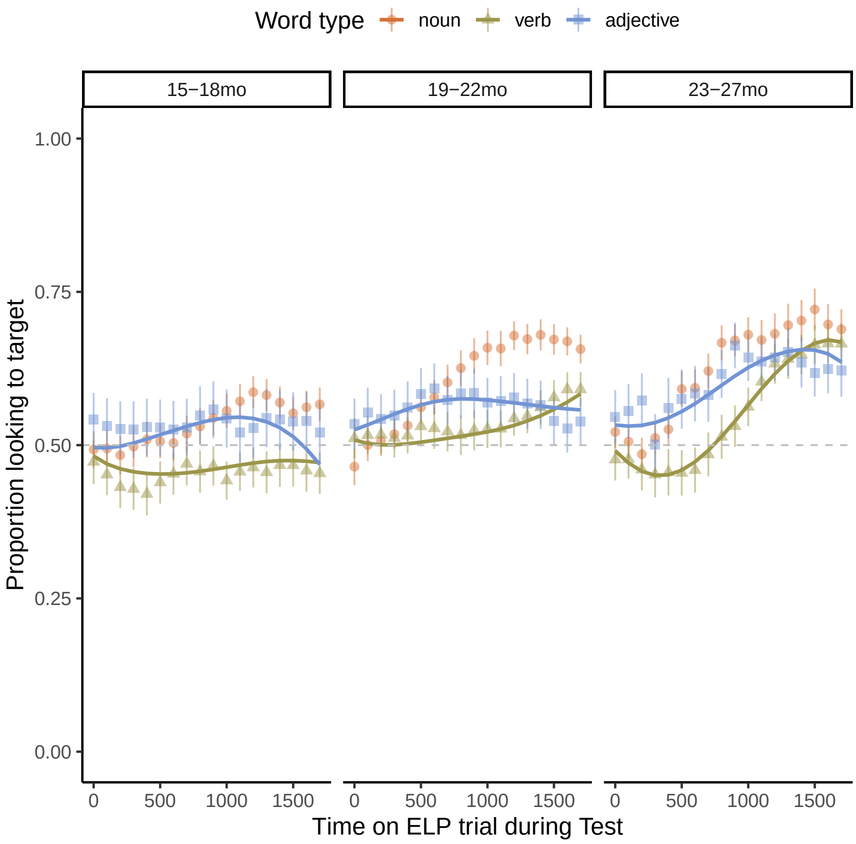


**Figure 2.11:** Model predicted proportion looking to target in verbs and adjectives from Comprehension trials including all difficulties by Time by Age during test. Raw looks towards nouns are also included for visual comparison although they were not part of this model. Grey dashed line depicts chance performance (0.50). Age in months was split in three age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

We explored correlations among word types and difficulties on ELP Comprehension (see Figure 2.12). We did not find significant correlations among different word types and difficulties using our threshold (i.e., sig.level <0.01). However, we see correlations between word types and difficulties and Comprehension which is expected since Comprehension includes the overall proportion of all those words.

We also looked at the relationships between ELP Comprehension and other

ELP measures. As can be seen in Figure 2.12, the correlation analysis shows positive relationships between overall proportion of looks to target in Reinforcement trials and Comprehension trials that include Moderate Nouns (t = 2.769; r = 0.347; p = 0.007). This means that children who are good at Reinforcement trials are also good at Comprehension trials that include Moderate Nouns.

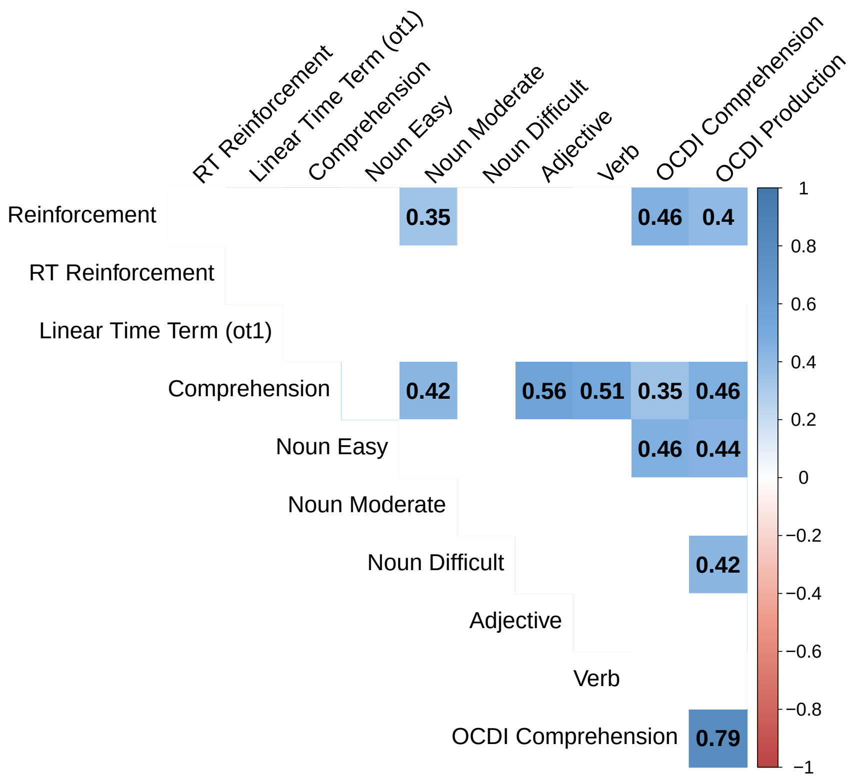


**Figure 2.12:** Correlation matrix for Reinforcement and Comprehension measures split by Nouns by difficulty, Adjectives and Verbs) and OCDI (Comprehension and Production total score). Only correlations smaller than 0.01 (sig.level <0.01) are displayed.Correlations were performed only on those children who had data for all the subtasks in this section. Positive correlations are in blue and negative correlations in red, the strength of the colour indicates the strength of the relationship.

A final key question is whether looks towards the target on ELP comprehension trials were related to parental responses on word comprehension checklists, in this case the OCDI. If ELP comprehension trials capture children's knowledge of words and parental responses are accurate, we would expect children to look more towards words that their parents said they knew. As can be seen in the correlation plot on Figure 2.12, we found a positive relationship between ELP Comprehension and OCDI Comprehension total scores ($t = 2.839$; $r = 0.354$; $p = 0.006$) as well as ELP Comprehension and OCDI Production total scores ($t = 3.877$; $r = 0.460$; $p = 0.000$; see Appendix for a plot showing the relationships between OCDI Scores and ELP Comprehension Figure A.7). These relationships are still significant when considering only easy nouns in relation to both OCDI measures as well as for difficult nouns and OCDI Production (for more details regarding these correlations see Figure A.6 in the Appendix).

**ELP Novelty Bias**

We used the familiarisation phase of Referent Selection trials to measure children's attraction to novelty. Novel images were paired with two types of familiar images: a) familiar images that appeared in Reinforcement trials, and b) familiar images that appeared in Comprehension trials. Both sets were familiar (easy) nouns, although nouns from Reinforcement trials were highly familiar (i.e., they have the highest frequency counts). We called this variable "Familiarity Image".

Proportion of looks to the novel image during Referent Selection familiarisation were fit with Time, Age and Familiarity Image. The binomial model included Time and participant interacting with Image in the random structure. The model showed a main effect of Familiarity Image and a 2-way interaction between Age and Familiarity Image (see F values on Table 2.12).

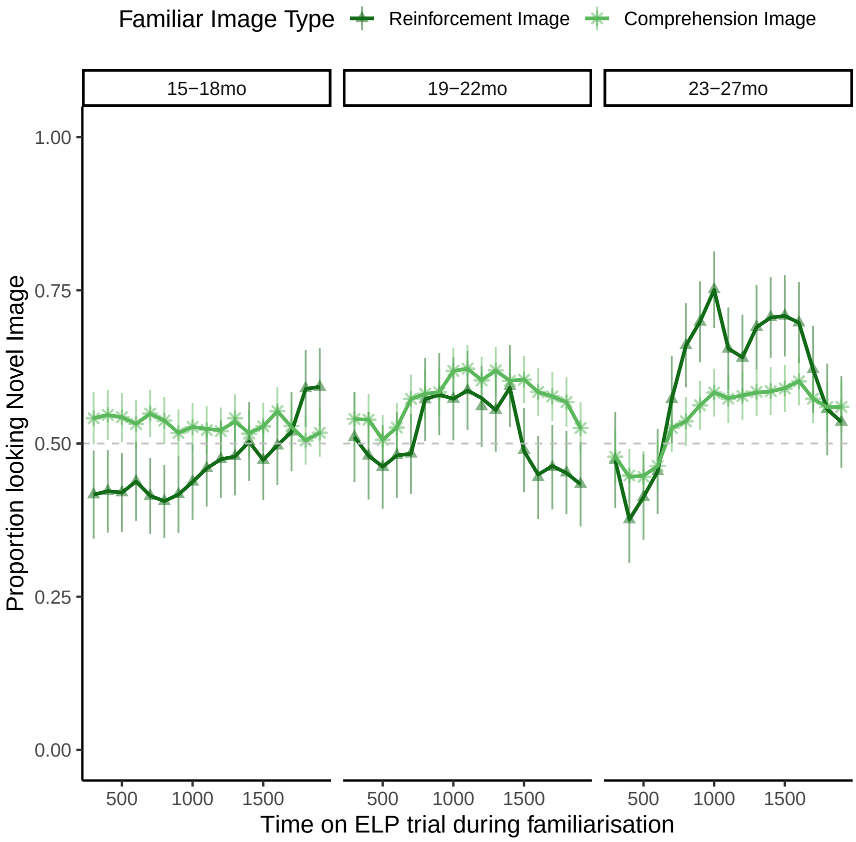


**Figure 2.13:** Model predicted proportion looking to the novel image in the familiarisation phase of Referent Selection trials. The model includes proportion looks to novel by Time by Age by Familiarity Image (a familiar image that appeared in Comprehension versus in Reinforcement trials). Grey dashed line depicts chance performance (0.50). Age in months was split in three age groups to facilitate visualization. Different shapes with standard deviation show the raw mean data per each 100 ms time bin split by Familiarity Image in different shades of green. Lines show the model predictions.

As can be seen in Figure 2.13, young children tend to be at chance level, with a slight preference for the familiar object that appeared in Reinforcement trials. By contrast, older children look more towards the novel object, particularly when the familiar image is highly familiar (i.e., also present on Reinforcement trials).

### ELP Referent Selection

To measure children's referent selection abilities, a binomial model mimicking the structure of the models used in the other ELP measures, assessed looks to target during the test phase of Referent Selection trials.

The binomial model included fixed effects of Time by Age by Word Type (novel versus familiar). In some Referent Selection trials the target was a novel object while in others it was a familiar noun. We accounted for that in the model. This model included Time and Participant interacting with Word in the random effects structure. Results from the model showed a main effect of the quadratic Time term and Age in months, a 2-way interaction between the quadratic Time term and Age in months, a 2-way interaction between the quadratic Time term and Word Type (novel versus noun), and a 3-way interaction between the quadratic Time term, Age and Word Type (see F Table 2.13).

Figure 2.14 shows that children look at the target in all types of referent selection trials, that is, when the target is a novel object-word and when it is a familiar one. Looking patterns show an age effect, with older children looking more and more consistently to the target. We do not see a main effect of word type (noun versus novel), which means children look overall to the target regardless. However, the rate of looking towards the target changes as the trial unfolds as a function of age and target type. These looking differences might be related to the novelty bias we observe in older children as those children tend to start the test phase already looking to the novel image before hearing the target word (i.e., they quickly look from the girl to the novel image prior to word onset).

Next, we explored relationships between novelty and familiarity biases and other ELP measures. The relationships that remained significant after correcting for multiple comparisons (sig.level <0.01) can be seen in Figure 2.15

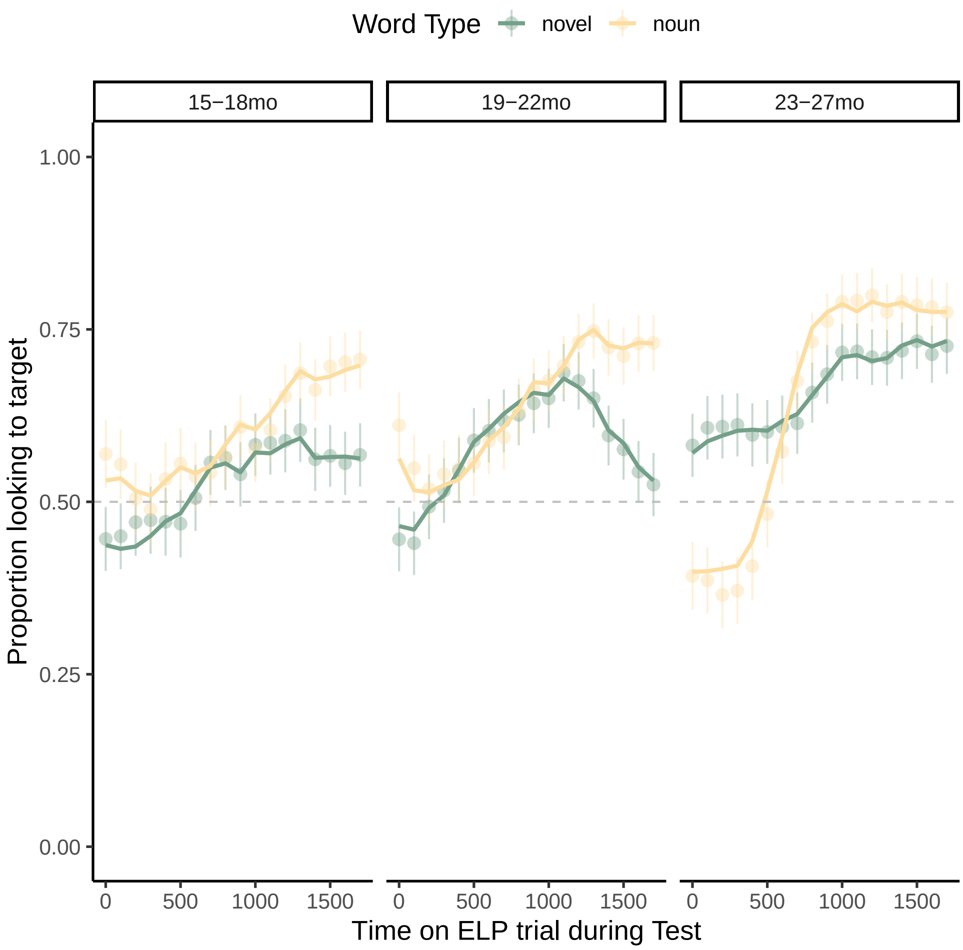


**Figure 2.14:** Model predicted proportion looking to the target in the test phase of referent selection trials. The model includes proportion looks to target by time in trial by age in months by word type (familiar versus novel). Grey dashed line depicts chance performance (0.50). Age in months was split in three age groups to facilitate visualization. Different colours show word type, in yellow looks to the target when that was familiar noun-object, in green looks to the target when that was novel word-object. Dots indicate the raw mean data per each 100 ms time bin including standard deviation. Lines show the model predictions

(see Appendix for more detailed information about the correlation coefficients, Figure A.8. Looks towards the novel object - novelty biases - were negatively associated with looks towards the familiar object - familiarity biases-. This makes sense because children that looked less at the novel object, implicitly looked more at the familiar object. Reinforcement trials were positively correlated with Referent Selection trials where the target was a familiar word (t

= 3.370; r = 0.383; p = 0.001). This significant correlation indicates that children who showed a stronger preference for the familiar image before being prompted with the target word, were also better at recognising a familiar target on Referent Selection trials when they were asked to do so (i.e., the target was the familiar and not the novel object).

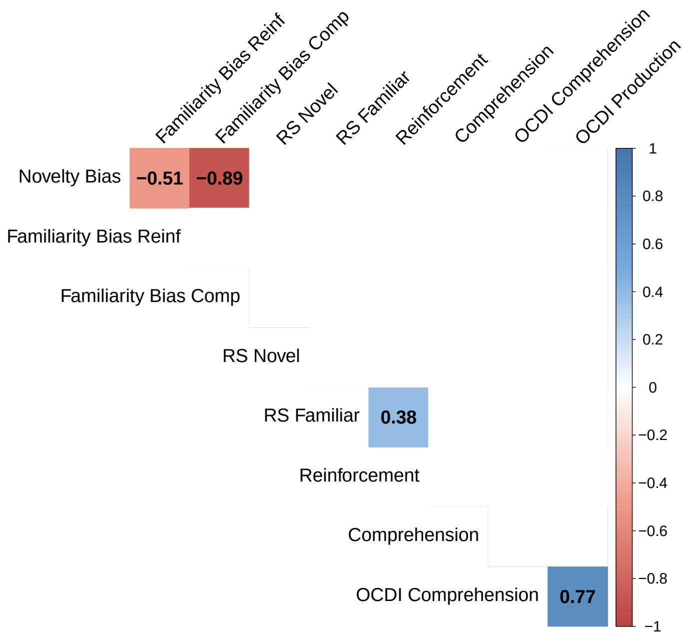


**Figure 2.15:** Correlation matrix for ELP measures including Novelty Bias, Familiarity Bias (for Reinforcement images versus Comprehension images) Referent Selection (RS) when the target was novel versus when it was familiar, Reinforcement, Comprehension and OCDI scores. Only correlations smaller than 0.01 (sig.level <0.01) are displayed. Correlations were performed only on those children who had data for all the subtasks in this section. Positive correlations are in blue and negative correlations in red, the strength of the colour indicates the strength of the relationship.

### ELP Retention

Following the GCA approach, a binomial model was fitted to the familiarisation looking data from the retention trials. Results from the model showed a significant interaction between the linear Time term and Age (see F Table 2.14). As can be see in Figure 2.16 children's looking responses are generally around chance levels, which means they do not show a preference for one image over another. However, the interaction we see in the GCA model indicates that children's looking patterns below and above chance vary as a function of age. In this sense, we can see a tendency to look to the "future" target later on in the trial in younger children, and a tendency to start the trial looking at the "future" target and switch to the distractor for older children.

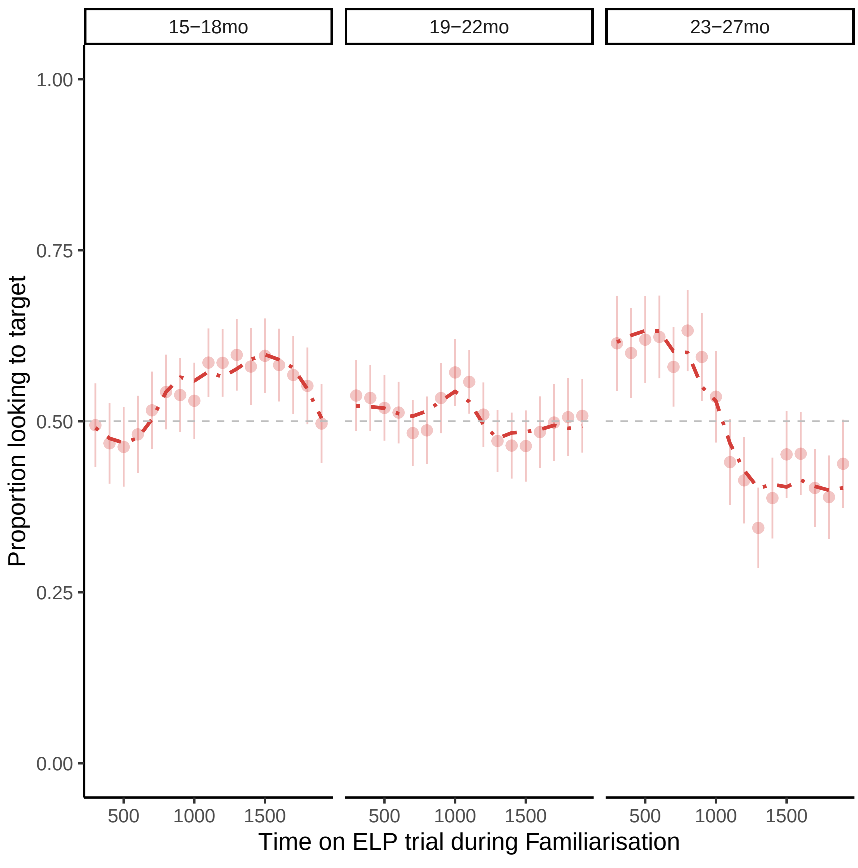


**Figure 2.16:** Model predicted proportion looking to target in Retention trials by Time and Age during familiarisation. Grey dashed line depicts chance performance (0.50). Age in months was split in three age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model prediction.

When modeling looking to the target during Retention test phase, we wanted to make sure children were paying attention and ”learning” the novel object-word mapping from Reference Selection trials, so we could assess whether they remembered the association after a 5 min break. For that purpose, we excluded Reference Selection trials in which the novel object was the target and participants looked at it less than 50% of the time (during the interest window).

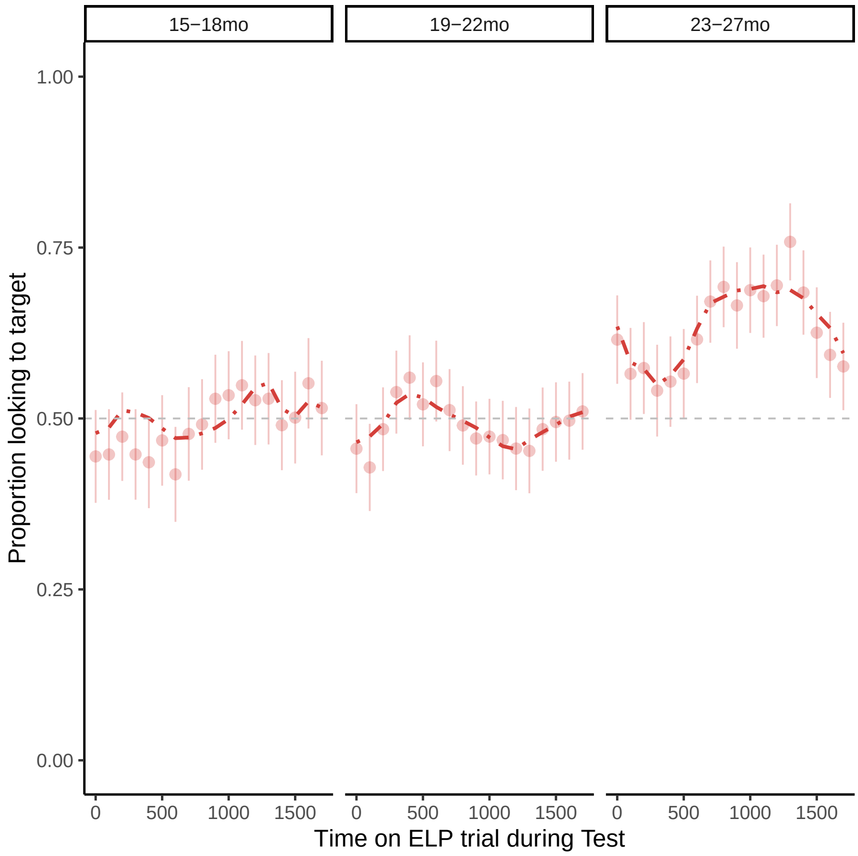


**Figure 2.17:** Model predicted proportion looking to target in retention trials by time and age during test. Grey dashed line depicts chance performance (0.50). Age in months was split in three age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model prediction.

Using the GCA approach, a binomial model was fit to the looking data from the test phase of the Retention trials. Results showed an interaction between the cubic Time term and Age (see F Table 2.15). As can be seen in Figure 2.17, young children do not show a preference towards the target nor the distractor after hearing the target word. However, older children look more towards the target by the end of the window of interest. This indicates that some older children are able to retain the previous word-object mappings.

In the correlation analysis, we explored relationships between ELP Retention and other ELP and vocabulary measures such as the OCDI. We see strong and positive relationships between Retention and Reinforcement trials ($t = 2.900$; $r = 0.358$; $p = 0.005$), as well as between Retention and Comprehension ELP trials ($t = 15.624$; $r = 0.898$; $p<0.001$), a relationship that holds when splitting by word type and difficulty. This indicates that children with higher proportions of looks to the target either in Reinforcement or in Comprehension trials, also show high proportions of looks to the target in Retention trials. It is possible that a better knowledge of highly familiar words helps learning and remembering novel word-object associations. The strong correlation between Retention and ELP Comprehension indicates that children who are better at identifying words from different types and difficulties (see the positive significant correlations between retention and easy nouns, moderate nouns, adjectives and verbs in Figure 2.18), are also better at learning new words. This might be because those children have larger vocabularies, which has been reported to help children learn new word-object associations (Kucker & Samuelson, 2012). In fact, we also see positive relationships between Retention and OCDI but only OCDI Production is significant ($t = 2.938$; $r = 0.359$; $p = 0.004$). This indicates that the ability to remember new word-object associations is related to production abilities more generally. Interestingly, we also see a positive relationship between Retention and RS Familiar trials ($t = 3.242$; $r = 0.391$; $p = 0.001$), in which the child has to select a familiar target in the context of a novel one. This means that children that are better at selecting a familiar target in the context of a novel one without the interference of novelty biases, are also good at remembering newly learned word-object associations. This is in line with the idea that children with larger comprehension abilities and larger (productive) vocabularies are better at selecting the correct object in the ELP task, which might help them remember new word-object associations (see Figure A.9 in Appendix for more details regarding correlation values.

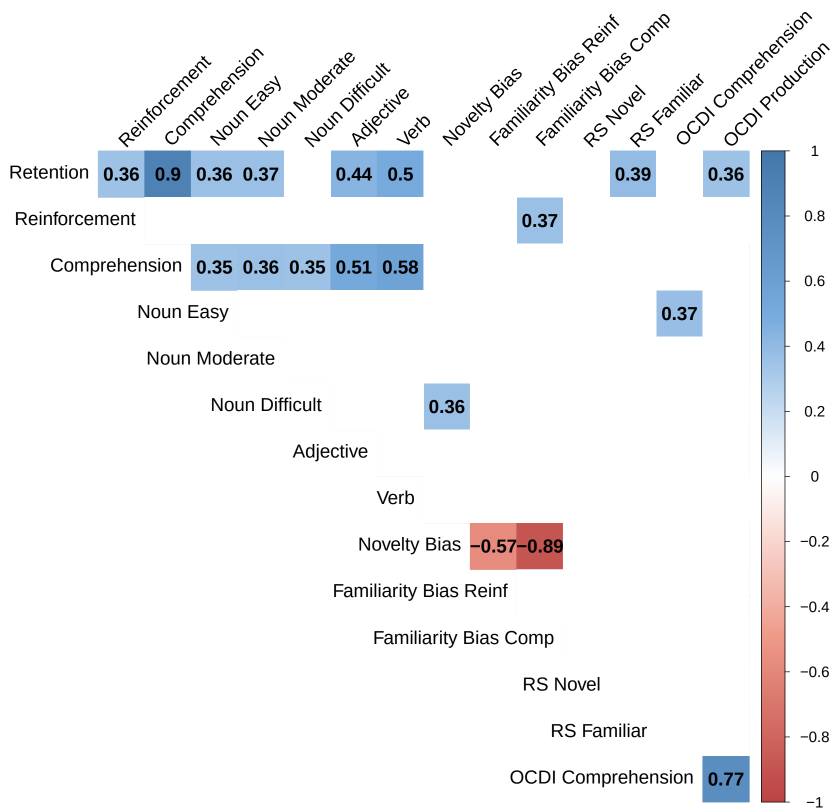


**Figure 2.18:** Correlation matrix for Retention and the rest of ELP measures including OCDI Comprehension and Production. Only correlations smaller than 0.01 (sig.level <0.01) are displayed. Correlations were performed only on those children who had data for all the subtasks in this section. Positive correlations are in blue and negative correlations in red, the strength of the colour indicates the strength of the relationship.

### 2.3.2 ELP Retest at 30 months

As in the 18-month-old group, we used a GCA approach on the 30-months-old retest data including all types of ELP trials (n = 76, 35 girls, between 28 and 36 months of age). This first model aimed to explore overall group patterns across trials before going into more detailed analysis and Test-Retest comparisons. A binomial mixed effects model was fit to the looking data extracted as proportion looking to target by Time in trial, by Trial Type, and by Age in months.

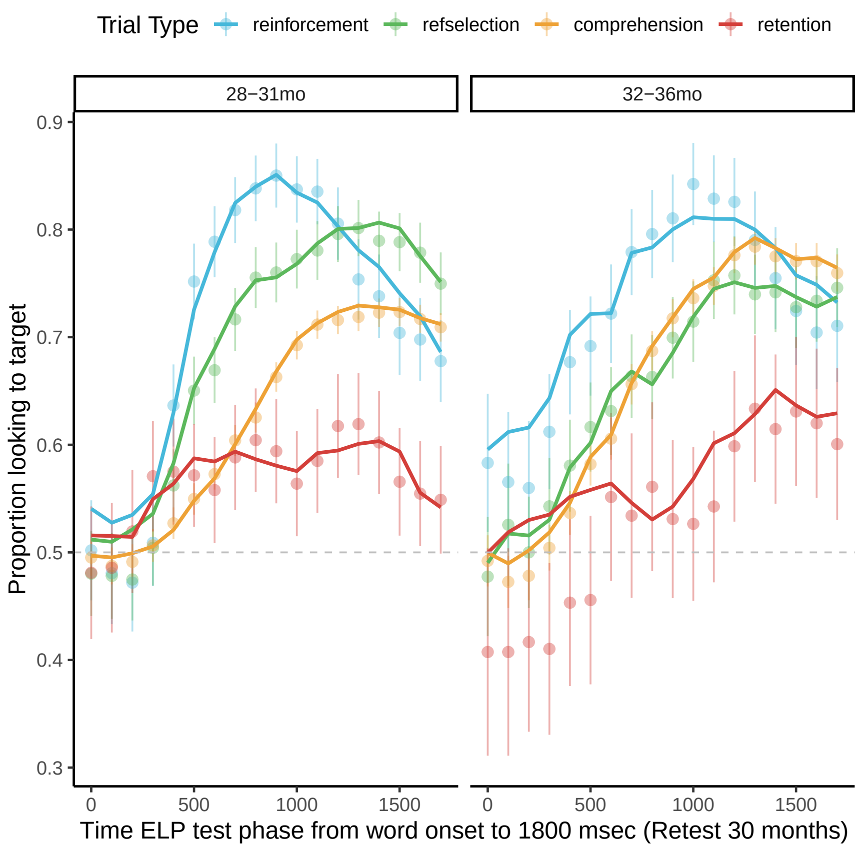


**Figure 2.19:** Model predicted proportion looking to target by Time by Trial Type by Age, during the window of interest of ELP test phase. Grey dashed line depicts chance performance (0.50). Age in months was split in two age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Lines in different colour show the model predictions by trial type.

The results from the model show similar main effects and interactions as those found for ELP Test at 18-months (see F Table 2.16) thus, results show similar group patterns to those seen at ELP Test. The most salient results are the strong age effects, particularly on Comprehension and, particularly, Retention trials. At ELP Test at 18 months, overall results showed that children were mostly at chance levels on Comprehension and Retention trials (except older children who showed slightly more looks to target by the end of the window of interest). As can be see in Figure 2.19, 30-months-olds' data show that all children look more towards the target with even higher rates at older ages.

### 2.3.3 Relationships between Test-Retest

This section addresses our two last questions: Q3) for those children who did ELP at both time points (Test and Retest), how did ELP performance at 18 months relate to performance at 30 months at the group level; Q4) and how did ELP performance at Test relate to performance at Retest at the individual level.

Note that we only compared data for three types of trials: Reinforcement, Comprehension and Retention. This was for two reasons. First, we wanted to measure if at 30 months, children had better retention abilities than at 18 months. Second, we wanted to measure the reliability of our new test by comparing performance across the Reinforcement and Comprehension.

#### ELP Reinforcement

At the group level, a GCA comparing test and retest shows several significant interactions (see Table 2.17). The most relevant to our questions are 4-way interactions between the linear, quadratic and cubic Time terms and Age, Reinforcement Pair Count, and Test Type (test, retest). These relationships indicate strong age affects related to the repetition of the word/image pairs. Younger children look at the target across the window of interest, while older children quickly look at the target (see the steepness of the curve in Figure 2.20), and then they release fixation to look to the other image. This suggests faster visual processing speed with the older children and more pronounced word repetition effects (recall that the Reinforcement word-image pairs repeated five times over ELP).

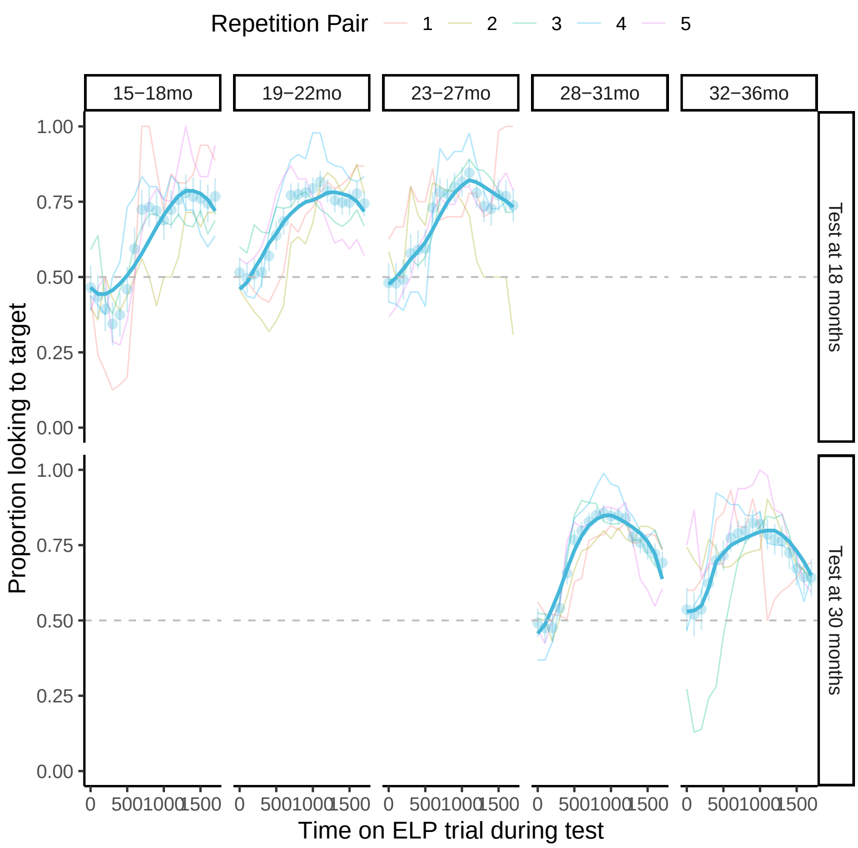


**Figure 2.20:** Model predicted proportion looking to target in Reinforcement trials by Age by Repetition Pair Count by Test Type (18 vs 30-months-old), during test. Grey dashed line depicts chance performance (0.50). Age in months was split in three age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions. Lines in different colours show repetition count.

At the individual level, we used overall proportion looking to target in the window of interest to assess if performance on Reinforcement trials at 18 months was related to performance at 30 months at the individual level. In particular, we wanted to examine if children with high proportions of looking to targets at 18 months also showed the same pattern when they were 30 months of age. A correlation analysis showed non significant correlations between performance at 18 months (test) and performance at 30 months (retest) for Reinforcement trials ($t = 0.939$; $r = 0.086$; $p = 0.349$). Similar results were obtained when using the model coefficients based on the linear time term ($t = 0.132$; $r = 0.011$; $p = 0.895$). This might be due to possible ceiling effects, because all children are very good at the highly familiar words presented on Reinforcement trials.

**ELP Comprehension**

A GCA was used to directly compare looks towards the target on overall Comprehension trials at both testing periods (18 and 30 months). Results from the binomial model showed a main effect of the linear and the cubic Time term, and a main effect of Age. We can also see 2-way interactions between the linear and the quadratic Time terms with Age (see F Table 2.18). These results indicate that older children look more and quicker to the target, as can be seen in Figure 2.21. As with Reinforcement trials, therefore, we continue to see age-related improvements in ELP task performance on Comprehension trials out to 36 months of age.

To assess whether individual children had similar looking patterns between test and retest, we used correlation analysis. We have already seen in the group analysis that children at 30-months look more to the target overall, which indicates that they are able to recognise the target word more reliably. However, that does not tell us if individual children showed a higher proportion of looking to the target during the retest period. It is worth noting that we tested the same children at both time points on the same words (with

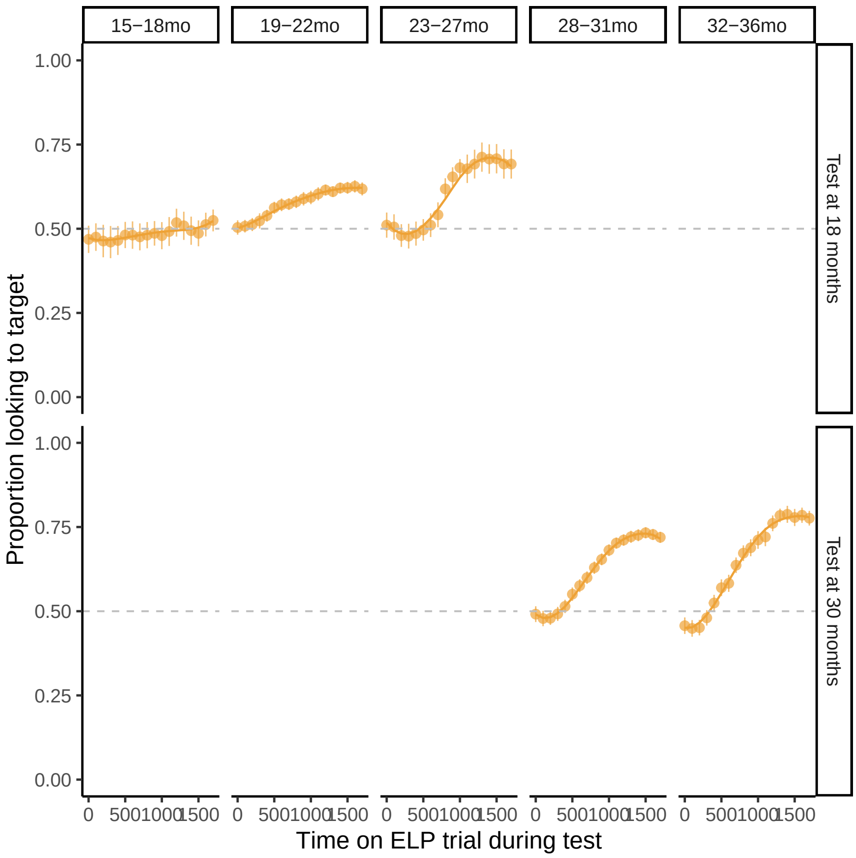


**Figure 2.21:** Model predicted proportion looking to target in overall Comprehension trials by Age by Test Type (18 vs 30-months-old), during test. Grey dashed line depicts chance performance (0.50). Age in months is split in groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

different orders). Thus, a possible outcome is that, since children have been exposed to those words more often, and they probably know them very well, children who showed high proportions of looks to target at 18 months should also show higher proportions of looking to the target at 30-months. To look into this possibility, we measured the relationship between mean proportion of looks to target at 18 months compared to mean proportion of looks to the target at 30 months (within the window of interest) for all Comprehension trials aggregated. We found a significant positive correlation between overall proportion of looks to the target on Comprehension trials at 18 months and overall proportion of looks to the target on Comprehension trials at 30 months ($t = 2.688$; $r = 0.316$; $p = 0.009$). Thus, those children who correctly identified many words at 18 months also identified many words at 30 months. We ran the same correlation analysis using model coefficients instead of proportions, in particular, the linear time term. This index reflects the rate of change of looks toward the target across trial time. A correlation comparing rate of change at 18 months versus at 30 months showed a positive relationship but it was not significant ($t = 1.858$; $r = 0.224$; $p = 0.067$).

The next step in our analysis examined Comprehension performance at the group level focusing on nouns split by difficulty. Proportions of looks to target on Nouns were fit with Time in trial, Word Difficulty, Age in months and test type (test, retest) as predictors. The random effects structure included Time by participant interacting with Test Type. The model results showed several significant effects with strong main effects of Time in trial, Age, Test Type and Word Difficulty (see Table 2.19). The most interesting relationship we see is a 4-way interaction between the linear, quadratic and cubic Time terms, Test type, Age and Word Difficulty. These results can be seen in Figure 2.22. The most striking finding is the strong developmental effect that is evident. Children between 15 to 18-months are at chance on all nouns, which means that they look equality at the target and at the distractor. With age, we start to see that most children can correctly identify the target when this is an easy or a moderate noun. Already at 23 to 27 months, some children can correctly identify the target when prompted with a difficult noun. Importantly, looks towards difficult nouns increases by 28 to 36 months, with children looking more often and much faster to words of all nouns difficulties. These differences across word difficulties indicate that ELP Comprehension is able to capture developmental changes in word learning. The looking trajectories also reflect a better knowledge of those words at older ages.

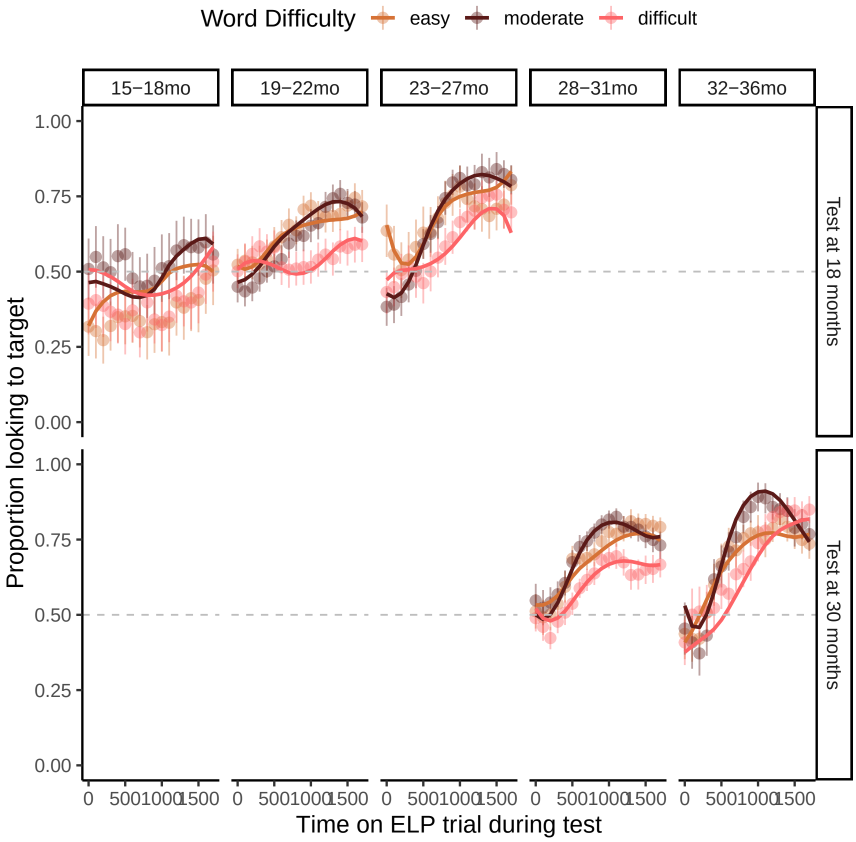


**Figure 2.22:** Model predicted proportion looking to target in Nouns on ELP Comprehension trials by Age by Word Difficulty by Test Type (test versus retest). Grey dashed line depicts chance performance (0.50). Age in months is split in groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

We also examined whether individual children performed in a similar manner across test and retest when considering only nouns. Overall proportions of looking to the target for nouns did not show a significant correlation between individuals across both observations ($t = 1.5632$; $r = 0.119$; $p = 0.119$). However, when splitting proportions by word difficulty, we find a significant positive correlation for difficult nouns ($t = 2.0797$; $r = 0.293$; $p = 0.043$). This finding might reflect that children perform quite good at easy and moderate nouns at test and retest and, therefore, performance is at ceiling for some children. With greater variation for difficult nouns, this measure effectively captures individual differences at test and retest. We also found a positive significant correlation when using the model coefficients (ot1). In this case, the linear time term at test and retest were correlated ($t = 2.048$; $r = 0.246$; $p = 0.044$). This relationship indicates that children who had faster rates of looking to the target over the course of the trial at test, showed faster rates at retest as well. This result arises when taking into account nouns overall (collapsing across difficulties). When splitting by noun difficulty, however, the linear relationship did not hold for any individual difficulty.

In a final set of analyses on Comprehension performance at test and retest, we conducted a GCA looking at adjectives versus verbs collapsing across word difficulties. Proportions of looks to target on Comprehension Adjectives and Verbs were fit with Time in trial up to the fourth term, Word Type, Age in months and Test Type (test, retest) as predictors. Time and participant interacting with Test Type were set as random effects. The model results showed main effects of the linear Time term as well as Word Type, there were several interactions between Time in trial, Age, Word Type and Test Type that can be seen on Table 2.20. These results indicate a similar pattern relative to what we found for Nouns only. As can be seen in Figure 2.23, most children start at chance levels for adjectives and verbs but, as they age, they show more looks to the target. Particularly on trials were the target was an adjective. By 23 to 27 months, some children are able to correctly identify the target when this is a verb. By 32-36 months, we can see that children are equally good for both adjectives and verbs, looking overall more to the target than at younger ages and getting there more quickly. Interestingly, the older children in our sample seem to be as good at verbs as they are at nouns by the end of our interest window.

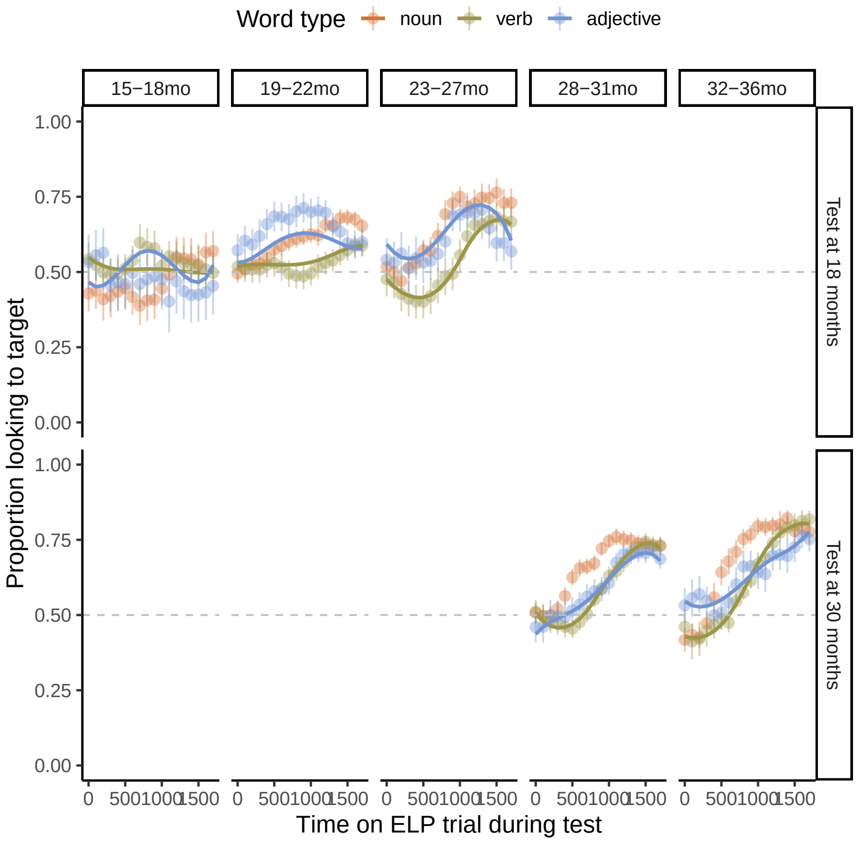


**Figure 2.23:** Model predicted proportion looking to target in Nouns on ELP Comprehension trials by Age by Word Type by Test Type (test versus retest). Grey dashed line depicts chance performance (0.50). Age in months is split in groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

As a final step in our analysis of Comprehension performance, we conducted a correlation analysis using overall proportion of looks to the target for verbs and for adjectives. We did not find a significant relationship between test and retest performance when aggregating across verbs and adjectives ($t = 1.081$; $r = 0.095$; $p = 0.281$). This relationship was not significant either when splitting between adjectives ($t = 0.135$; $r = 0.017$; $p = 0.892$) and verbs ($t = 1.596$; $r = 0.197$; $p = 0.115$) even though it showed a positive trend for verbs. We also explored relationships between test and retest using model coefficients, that is, the linear time terms instead of mean proportions. Again, we did not find significant correlations when aggregating across adjectives and verbs (t = 0.457; r = 0.057; p = 0.648). Neither for adjectives (t = 0.074; r = 0.009; p = 0.94) alone nor verbs (t = 1.963; r = 0.240; p = 0.054). For an interested reader, a plot containing all the correlations reported in this section looking at ELP Test Retest Comprehension can be found in the Appendix, Figure A.10.

### ELP Retention

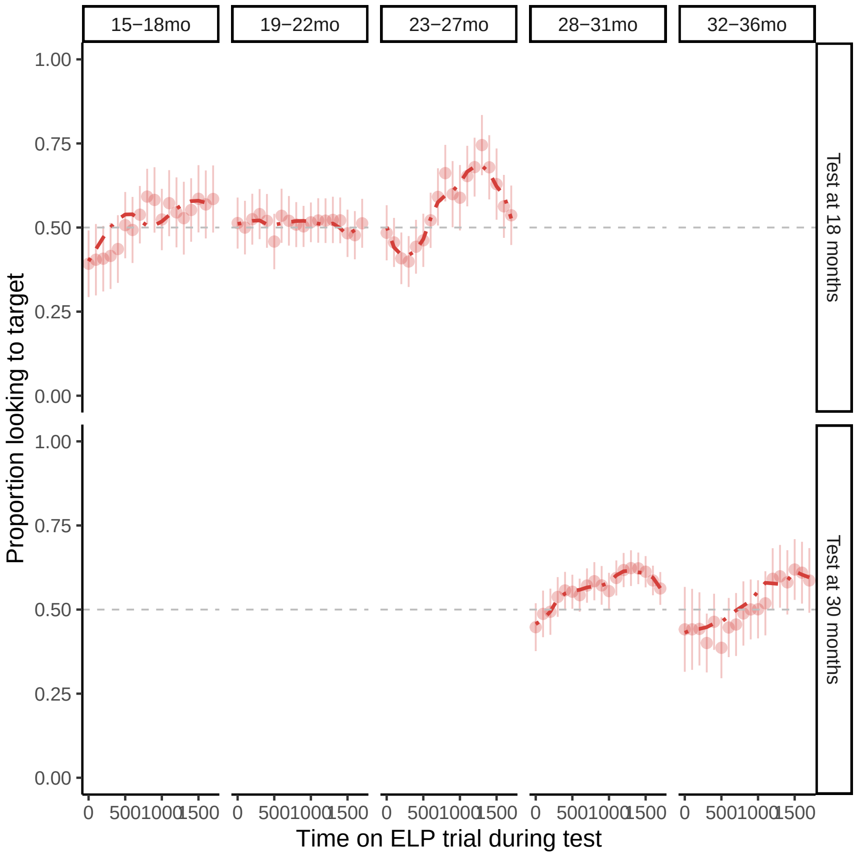


**Figure 2.24:** Model predicted proportion looking to target in ELP Retention trials by Time by Age by Test Type (Test vs Retest). Grey dashed line depicts chance performance (0.50). Age in months is split in groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

A GCA using a binomial model with Time to the third order and Age and

Test Type, predicting looks to target on Reinforcement trials showed a main effect of the cubic Time term and a 2-way interaction between the cubic Time term and Age (see F Table 2.21). As can be seen in Figure 2.24, only children between 23 - 27 months of age show robust retention, looking more to the target by the end of the window of interest. Some older children seem to look at the target, also by the end of the window of interest, but overall children did not show a strong preference to look at the target during the retest.

At the individual level, we found a significant positive correlation between retention during test and retest ($t = 2.754$; $r = 0.387$; $p = 0.008$). This relationship is significant when using both proportion looking to target and model coefficients (the quadratic time term, or ot2, in this case; $t = 1.980$; $r = 0.289$; $p = 0.054$).

## 2.4 Discussion

In this study we created a new measure of early language abilities, the ELP task, which integrates language processing measures of speed of word processing, online word comprehension, novelty bias, referent selection and retention of new words. We collected data from a large number of children between 16 and 27 months of age on five different language processing measures. This data was used to examine the relationships between performance across different ELP measures at both the group and individual level, and look at changes across age in our measures. We also collected retest data from a subset of the same children at a second time point between 30 and 32 months of age. The data allowed examination of retention abilities in older children which should improve based on previous findings (Bion et al., 2013), as well as an assessment of the reliability of the ELP measures across two time points. Our results indicate that the ELP task is an effective, direct measure of early language processing, that captures multiple critical processes that support early vocabulary development, as well as developmental and individual differences in a 15 - 20 minutes assessment.

Moreover, ELP was created to be a direct and automatised measure of language processes. For that purpose, the task incorporates a remote eye-tracker, extracting measures of looking time and gaze trajectories as dependent measures at the millisecond level. This allowed us to implement fine grained growth curve analysis techniques in our analyses instead of only using overall proportions and reaction time, which we used to infer looking trajectories. This has been particularly useful when looking at large age ranges, were overall proportions do not give the details that we can extract from looking dynamics over the trials. Moreover, eye-tracking techniques allow testing young participants contributing to better understand early language processes at very young ages. Using automatic, easy-to-implement protocols has also allowed us to gather large participant samples than looking tasks that rely only on coding techniques, increasing the power of our statistical analysis.

We started this project with four questions: Q1) Can ELP measure different language processes? Q2) Can ELP be used to capture individual differences? Q3) Is ELP reliable as a tool and able to capture developmental trajectories between 18 and 30 months? Q4) Is ELP able to measure individual developmental trajectories between test and retest?

### 2.4.1 ELP as a measure of Language Processes and Individual Differences

Our first question (Q1) was if ELP could measure different language processes in a single task. Our results show that children's looking patterns are very different across ELP measures, indicating that the task is sensitive to different language processing abilities. Moreover, our results indicate that ELP is highly sensitive to developmental effects. All our analysis shows that older children are faster and better at ELP. Crucially, although all ELP measures showed maturity effects, those effects were different across measures, reflecting different language processing abilities that become more efficient as children's age. Finally, ELP proved to be sensitive to individual differences, answering our second question (Q2). Below we discuss our findings in relation to those first two initial questions in the context of previous literature.

**Speed of Word processing**

On ELP Reinforcement measures, younger children were slower, whereas older children could quickly identify the spoken word. This is in line with previous findings showing the speed of word processing increases with age, as children gain experience with language (A. Fernald et al., 1998, 2006). The fact that older children on ELP Reinforcement were very quick at identifying the correct images suggests that older children knew the words we used on reinforcement trials, which are highly familiar, better than younger children. Our results also show that children tend to look more reliably to the target as the number of repetitions increases, indicating that repetition helps performance. However, in older children we observe that after quickly identifying the target, they look away. It is possible that older children could be experiencing an habituation effect, which might explain why they look away quicker than younger children. The idea that faster looks to the target is related to children's lexical abilities is supported in our results showing that ELP Reinforcement and ELP Comprehension performance (moderate nouns) is correlated in individual children. This means that children who knew many words (i.e., high OCDI scores) and were good at online comprehension, were also good at recognising highly familiar words. This is expected, since the highly familiar words from our Reinforcement trials were likely in their part of their vocabularies. These results replicate previous findings showing that children with larger vocabularies are more efficient at word recognition (e.g., A. Fernald et al., 2006; Hurtado, Marchman, & Fernald, 2008b; Mahr & Edwards, 2018; Donnelly & Kidd, 2021). A difference here, however, is that in some of those studies vocabulary size was calculated based on parental report instead of online comprehension. We did not find relationships between ELP Reinforcement measures and OCDI scores. This could be for two reasons. First, vocabulary measures based on parental report at those ages might be less robust than online comprehension, particularly when measuring expressive vocabularies, since young children might not produce many words. This would make it harder to replicate similar results across studies. Second, we might not find the same relationships between parental report measures of vocabulary and our measures of speed of word processing because we used a different analysis method. A. Fernald et al. (2006) analysed each age group separately. Given the fact that they did not find significant relationships between parental report and speed of word processing measures – neither at 18, nor at 21 months of age – it is possible that we do not find any relationships between our speed of processing measure and OCDI because we collapsed across a large age range (15 to 27 months) in our correlation analyses. Posthoc analysis splitting ELP Reinforcement by age group could help determine if there are relationships between some of our age groups and OCDI comprehension and production. Nonetheless, we were able to measure children's online comprehension abilities using ELP Comprehension trials, which were highly correlated with OCDI scores.

Another reason why we see divergences between some of the previous findings and our speed of word processing measures is because they are slightly different. For example, our reaction time (RT) measure for speed of word processing on ELP Reinforcement was not predictive in opposition to what other studies have reported (e.g., A. Fernald et al., 2006). ELP featured a more sensitive measure, looks always started in the center (i.e., looking at the girl). We calculated the latency (RT) of the first look to the target after word onset from the center. We designed the task this way for two reasons. First, to be able to include most of the trials in our analysis. Second, to capture children's ability to process a word using partial phonetic information. However, it is possible is that this lead to a noisier measure particularly in younger children, which might have looked at the target (on that first look) accidentally. Literature has shown that children can process words based on partial phonetic information – adults and older children are quite good at it, even some 18-months-old children with larger vocabularies (A. Fernald et al., 2001). However, some of our children were younger than that, and it might have been harder for them to process a word from the beginning leading to some noisy looks to target and distractor. Our RT might have some of those looks. Another issue is that we added all our data in a single correlation analysis and, thus, possible age effects in processing abilities based on RT were not accounted for, neither vocabulary size. Future analysis taking age and vocabulary size into account could help to bring some light into that issue.

We did not find relationships either between accuracy and the linear time term of our model coefficients as reported in Mahr and Edwards (2018). However, in that study children were overall older (39–52 months) than in ours. Thus, it is possible that this effect is too small to be seen in a younger sample. For instance, in Figure 2.7, it is noticeable how similar the linear time terms are (i.e., the steepness of the curve) for 15 to 22 months old children. A model looking at each age group separately could help disambiguate, if this relationship is present in the older children of our sample.

**Online Comprehension**

In contrast to ELP Reinforcement trials, we found positive relationships between both OCDI comprehension and production total scores, and ELP Comprehension. These significant correlations with OCDI production and comprehension were highly significant for easy and significant between difficult nouns and OCDI production. This suggests that ELP Comprehension measures are related to parental responses on word comprehension checklists to some degree. Is is not surprising that we find strong relationships between easy nouns and OCDI responses. First, easy nouns might appear frequently in children's daily interactions, which means that parents might have a lot of opportunities to notice their children's understanding of those words. Second, parents might be better at reporting children's understanding of concrete words, such as nouns, because children's knowledge of those words or the mapping between word and object might be easier to notice then for more abstract words. It is less clear the relationship between difficult words and OCDI Production. A possibility is that parents are good at remembering the difficult words that their children produce (e.g., mosquito or crocodile) because they are more rare and "special".

The fact that ELP Comprehension is associated with OCDI is important because we based our task on the CCT task. The CCT selected their words based on the CDI, reporting high reliability with this measure (Friend & Keplinger, 2008). This indicates that our task is accurately measuring word comprehension abilities. Nevertheless, we slightly deviated from the CCT, because we selected and classified our words based on frequency of appearance in a naturalistic language corpus. The fact that we find relationships between online word recognition and OCDI scores in our task means that children's lexical knowledge is related to frequency of occurrence of those words in their home input, as previous literature has suggested (Goodman et al., 2008).

Results from ELP Comprehension, also showed a strong age effect, because older children looked more, earlier and faster to the target than younger children on those trials. This effect was consistent when splitting across word types and difficulties. As one could expect, children are better at recognising easy and moderate nouns that difficult nouns, looking more and more consistently to the target on easy and moderate noun trials. This indicates that children have more trouble recognising difficult nouns. This is consistent with previous literature using the CCT task (Friend & Keplinger, 2003). Children, particularly older ones, were also better at recognising adjectives than verbs. In fact, older children were as good as recognising adjectives as they were at recognising nouns. We see a switch in our data from 19 - 22 months to 23 - 27 months where children start to show evidence to be able to efficiently recognise adjectives. Since most adjectives describe properties of objects, we could imagine that children first need to have a good knowledge of nouns before the are able to effectively recognise adjectives – and this is why we see this shift (Booth & Waxman, 2009). Thus, it could be expected that English learning children learn nouns first, and then the adjectives related to those nouns. This idea is supported by our findings, showing that individual children who were good at identifying adjectives also have good overall comprehension abilities, because they are able to recognise many of the words that appear on ELP Comprehension. This means that they know words from different types and difficulties.

Verbs however, were challenging in our sample, because only older children showed looks greater than chance levels on trials with verbs, and mostly at the end of the window of interest. This might reflect fit with prior findings in literature suggesting that nouns are acquired before verbs because the concepts underlying most nouns are more concrete, or imaginable, than those underlying most verbs (e.g., see Waxman et al., 2013 for a review). Moreover, our data suggests that overall comprehension abilities could be related to knowing "hard words" such as verbs, because we see a relationship between Verbs and overall Comprehension on ELP. It is possible that children who are more experienced with language, know more different types of words including verbs. However, this could also reflect our use of 2-dimensional images to represent verbs. Younger children in particular might have less experience using 2-dimensional representations of actions, making it harder for them to identify their meaning. This difficulty in identifying verbs from 2-dimensional images might also be reflected on the delay we see at older ages, when children only look at the target by the end of the window of interest on verb trials, compared to noun and adjective trials. Still, it is possible that more experience with language helps children overcome this disadvantage.

It is surprising we do not see a relationship between easy nouns and reinforcement because both types of trials contain easy nouns. This could be due to a ceiling effect since all children perform overall quite well at both. However, moderate nouns show more variability in overall proportion of looks to target, which allows us to capture individual differences.

### Novelty Biases

Results from ELP Novelty biases showed that young children have a slight preference for the familiar object that appeared in Reinforcement trials, but were mostly at chance levels. By contrast, older children looked more towards the novel object, particularly when the familiar image was highly familiar (i.e., also present on Reinforcement trials). This looking pattern is surprising, as usually young children look more to the novel object (e.g., Kucker et al., 2018; Horst, Samuelson, Kucker, & McMurray, 2011). Note, however, that the ELP task differs from previous tasks in that children are asked to look towards images indicated by the word on every trial. Because referent selection trials occur later in the task, it is possible that we see a weaker novelty bias in ELP due to this pervasive looking to familiar items. Moreover, ELP used images instead of real objects, which might also play a role in the strong novelty biases in prior work where children can reach out and grab the novel objects. Furthermore, there is some indication in literature that at younger ages (i.e., around 15 months), children actually prefer familiar things. This is the case of accented speech. We know that 3-year-old children prefer native-accented speakers regardless of the speaker accuracy when naming an object. In contrast, 4-and 5-year-old children endorse the names provided by the accurate speaker, regardless of the accent (Corriveau, Kinzler, & Harris, 2013). Thus, it is possible that the role of novelty biases varies over development as a function of the context where it occurs.

**Referent Selection**

Results from ELP Referent Selection showed that children were able to correctly identify the target in all types of Referent Selection trials, that is, when the target was a novel word and when it was a familiar one, and that this ability increased and became more consistent with age. This is consistent with previous literature using 2-dimensional stimuli (Bion et al., 2013) and 3-dimensional objects (Horst & Samuelson, 2008). In our Referent Selection measure, children were very good at disambiguating (i.e., finding the novel object in response to the novel word) as well as matching the known word to the familiar object that appeared in the context of a novel one. However, the amount of looking towards the target changed over the course of the trial – particularly at older ages – with older children having an initial preference for novelty, regardless of the target. These looking dynamics might be related to the novelty bias we observe in older children before they hear the target word. This further supports the idea that older children might show a novelty bias in the context of highly familiar images in this task. Literature shows that novelty biases might help children in the disambiguation process, where they need to map a new word to new object, and create a new association (Kucker et al., 2018). Older children might be more resistant to the "familiarity priming" of ELP and use the novelty bias to learn the new word. In contrast, the familiar design of ELP might have helped children perform well on familiar words, in contexts where referent selection can be hard (i.e.,paired with a novel object). This could explain why we find a relationship between Reinforcement trials, containing highly familiar words, and referent selection of familiar targets.

**Retention of new words**

Results on ELP Retention measures show that children's looking responses are generally around chance levels in the younger children of our sample, which means they do not remember the new word-object mappings that were established on referent selection trials. This is in line with previous literature showing that young children are quite good at selecting the correct referent but that have trouble remembering newly learned word-object associations after a short delay (Horst & Samuelson, 2008; Kucker & Samuelson, 2012). However, our results also indicate that some of the older children look more towards the target by the end of the window of interest. These findings are very similar to Bion et al. (2013) reporting that 30 months old children performed slightly above chance levels in their study (*M* = 0.59; *sd* = 0.18, t = 2.11, p = .049 in comparisons against chance 0.50). Even though at the group level older children in both studies are not very robust, we find strong correlations between ELP measures and Reinforcement, indicating that some children are able to remember the new words-object associations. Our results show relationships between ELP Retention and ELP Reinforcement, ELP Comprehension (across word types and difficulties), OCDI production scores, as well as a relationship between Retention and Referent Selection in familiar trials. These relationships highlight the role of lexical knowledge in word learning. Individual children in our task who had good word recognition abilities, were also good at word learning. This is in line with literature highlighting the role of lexical knowledge and vocabulary size on word learning (Bion et al., 2013; Samuelson et al., 2017). Our findings suggest that children who were able to remember the word-object associations were those that also had better lexical skills. Thus, it is possible that a good knowledge of different types of words enhances learning and remembering novel word-object associations in our task. This suggests that the ability to remember new word-object associations is related to production abilities in general, but is in line with the idea that a stronger and larger lexical knowledge supports children's ability to remember newly learned words. Also, performance in ELP Retention trials was associated with Referent Selection trials where the target was familiar (in the context of a novel image). The presence of this positive relationship suggests that children were good at "discarding" the novel foil when the novel item was the distractor. We do not find a relationship between looks to the novel image during referent selection and retention of the novel word-object association. Given the fact that most of our children (particularly the younger ones) did not remember the new word, this result is not surprising. However, older children with high lexical skills – or with good disambiguation abilities – might be able to correctly map the new word into the correct referent and remember it later on. Similar results, regarding Retention abilities and vocabulary skills have been reported at the group level in 30-months-old children (Bion et al., 2013).

Overall, these effects in relation to the different ELP measures reflect how ELP is able to capture different patterns of language processing abilities in individual children. Moreover, at the group level, the task is sensitive to developmental change. Older children know more words, are faster processors, are skilled in using novelty biases to disambiguate between familiar and novel objects, are better at making new word-object mappings and they can also remember them.

### 2.4.2 ELP Relationships across Development: Test Retest

So far we have seen that the ELP task is able to capture group and individual differences across several language processing measures in data from children who span a large age range. The two last questions we had at the begging of this project refereed to children's language processing abilities across development. Our third question (Q3) investigated if ELP was able to capture developmental trajectories between test and retest. The final question of this project (Q4) looked into individual developmental trajectories between ELP test and retest.

Our ELP test-retest data confirmed that the task was accurate to capture multiple language processes and developmental effects over a large age range. At the group level, performance increased with age for both ELP Reinforcement and ELP Comprehension with older children showing better skills than younger ones.

ELP Reinforcement showed interesting effects related to the repetition of the word-image pairs. Younger children in our sample look at the target across the window of interest, while older children quickly look at the target and then they release fixation to look to the other image. This suggests faster visual processing speed in older children but also more pronounced word repetition effects. This replicates what we found in the 18 months group test, but with accentuated age affects. This indicates that older children might have a good knowledge of the highly familiar words and, thus, ELP is still able to capture speed of processing efficiency at older ages. However, older children might find Reinforcement words too easy and uninteresting by the end of the trial and, thus, they look away quicker than younger children. This finding brings new data regarding speed of word processing of highly familiar words from an age group that is not very well documented, adding to the evidence provided by Mahr and Edwards (2018).

ELP Comprehension test-retest data also showed age-related improvements in ELP task performance. This also becomes clear when splitting across word types and difficulties. Older children were very good at recognising nouns, verbs and adjectives. Children continued to improve with age in their ability to recognise words from different difficulties. This effect is particularly clear in difficult words, indicating that our categorisation of word difficulty based on frequency of occurrence in natural speech is in line with the learning pattern that would be expected across development. This suggests that word frequency is a good indicator of lexical development (Goodman et al., 2008), and adds further evidence that the task is stable over development, while still being able to capture differences across word recognition abilities. This is possible thanks to adding ”hard words” such as verbs and difficult words in our task. This is probably the reason why we did not see significant relationships between individuals on test-retest Reinforcement trials, which contain only highly familiar words. Our explanation is that most children might be at the ceiling, not allowing the task to measure individual differences. However, since ELP Comprehension had words varying in difficulty, we can see relationships between individual's performance at test versus at retest. Those children who correctly identified many words at test when they were approximately 18 months, also identified many words at retest when they were 30 months, further adding to the evidence of the stability of the task at the group level and at the individual level.

By doing a retest at 30 months, we were also able to measure if retention abilities improved with age. That would be expected based on previous studies (Bion et al., 2013) and theoretical arguments relating retention with better lexical skills. However, Retention was the only ELP measure were children did not become more skilled with age. What we see in our data is that only children between 23 - 27 months of age showed robust retention. Some older children seem to show some indication of retention but overall, children did not remember the new word. It is possible that these results are not robust because we have fewer trials and the data is noisier. That said, we found individual relationships between individual children across test and retest on Retention trials. This means that children's performance was consistent on this measure, and suggests that although most children are at chance level, Retention is still sensitive to individual difference at older ages.

These results answer Q3 and Q4 regarding the stability of ELP across measurements. Highlighting that ELP is can capture consistent differences on language processing abilities over development.

### 2.4.3 Limitations and Future Directions

It would be useful for future work to examine the relationships across the ELP measures in more depth. We used only basic correlations to look at the relationships between ELP measures. But we could extend those analysis to more complex models that include several ELP measures, as well as age, to see how they relate to each other in a more fine grained way, as well as assess how they might predict one another. Moreover, ELP gives us measures of how children use their language abilities in the moment. However, we do not know how their previous experiences (such as language input or SES) might influence their performance on the task. Thus, further research is needed to better understand the relationships between children's previous experiences and their later language abilities. Mediation analysis or analysis of growth curve including contextual variables related to children's environment would contribute to our understanding of how previous language experiences influence word learning and processing skills (see Hurtado et al., 2008b and Mahr & Edwards, 2018 for two different examples of statistical analysis relating language input, language processes and vocabulary size).

Another question is how ELP might generalise. Most of our participants came from high and middle income households and had highly educated parents. Data collection should extend to a more heterogeneous populations to be able to assess how our results might apply to children from different cultural backgrounds.

## 2.5 Conclusion

Studying the processes that support early word learning is crucial to better understand how they influence later language and cognitive skill. Measuring individual differences early in development can help to better understand how those differences emerge, and how they influence language development. However, while variation in early language development is a key predictor of later cognitive abilities, many of the studies in this literature use indirect measures of language development such as parental report of words understood and produced, or tasks that measure only a particular language skill.

In this study, we created a new measure of early language ability. The ELP task integrates measures of speed of word processing, online word comprehension, novelty biases, referent selection and retention of new words. Data from this task highlight that the ELP is an effective and reliable measure of multiple early language processes, that can capture individual and age differences over development. The next step is to look at the relationships between language processes measured using ELP, and children's previous language experiences.

These results set the stage for future work to measure language processes in infancy in order to predict longer-term language and cognitive outcomes, as well as working to understand how early language processing abilities lead to differences in language skill over development. Importantly, understanding the mechanisms that underlie these relationships could provide empirical evidence that inform intervention efforts early in development.

## 2.6 Significance Tables

**Table 2.5:** Regression results for ELP at 18 months across all trials during Test

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 334.90 | 1.00 | 0.00 | *** |
| ot1 | 15.26 | 1.00 | 0.00 | *** |
| ot2 | 14.25 | 1.00 | 0.00 | *** |
| ot3 | 0.53 | 1.00 | 0.47 | |
| Age | 18.23 | 1.00 | 0.00 | *** |
| TrialType | 975.54 | 3.00 | 0.00 | *** |
| ot1:Age | 8.23 | 1.00 | 0.00 | ** |
| ot2:Age | 8.07 | 1.00 | 0.01 | ** |
| ot3:Age | 2.67 | 1.00 | 0.10 | |
| ot1:TrialType | 170.83 | 3.00 | 0.00 | *** |
| ot2:TrialType | 152.51 | 3.00 | 0.00 | *** |
| ot3:TrialType | 74.64 | 3.00 | 0.00 | *** |
| Age:TrialType | 721.39 | 3.00 | 0.00 | *** |
| ot1:Age:TrialType | 123.59 | 3.00 | 0.00 | *** |
| ot2:Age:TrialType | 103.23 | 3.00 | 0.00 | *** |
| ot3:Age:TrialType | 68.20 | 3.00 | 0.00 | *** |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months and Trial Type including all ELP trials. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.6:** Regression results for Reinforcement trials at 18 months during Familiarisation

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 0.01 | 1.00 | 0.92 | |
| ot1 | 1.12 | 1.00 | 0.29 | |
| ot2 | 5.02 | 1.00 | 0.03 | * |
| ot3 | 0.29 | 1.00 | 0.59 | |
| Age | 0.10 | 1.00 | 0.75 | |
| RepetitionPairCount | 0.04 | 1.00 | 0.84 | |
| ot1:Age | 0.56 | 1.00 | 0.45 | |
| ot2:Age | 4.24 | 1.00 | 0.04 | * |
| ot3:Age | 0.44 | 1.00 | 0.51 | |
| ot1:RepetitionPairCount | 3.46 | 1.00 | 0.06 | . |
| ot2:RepetitionPairCount | 0.02 | 1.00 | 0.90 | |
| ot3:RepetitionPairCount | 0.00 | 1.00 | 0.97 | |
| Age:RepetitionPairCount | 0.08 | 1.00 | 0.78 | |
| ot1:Age:RepetitionPairCount | 0.30 | 1.00 | 0.58 | |
| ot2:Age:RepetitionPairCount | 0.14 | 1.00 | 0.71 | |
| ot3:Age:RepetitionPairCount | 0.10 | 1.00 | 0.75 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months and Repetition Pair Count. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.7:** Regression results for Reinforcement trials at 18 months during Test

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 2.22 | 1.00 | 0.14 | |
| ot1 | 0.32 | 1.00 | 0.57 | |
| ot2 | 26.80 | 1.00 | 0.00 | *** |
| ot3 | 0.14 | 1.00 | 0.71 | |
| Age | 2.71 | 1.00 | 0.10 | . |
| RepetitionPairCount | 2.81 | 1.00 | 0.09 | . |
| ot1:Age | 4.00 | 1.00 | 0.05 | * |
| ot2:Age | 22.36 | 1.00 | 0.00 | *** |
| ot3:Age | 1.02 | 1.00 | 0.31 | |
| ot1:RepetitionPairCount | 2.56 | 1.00 | 0.11 | |
| ot2:RepetitionPairCount | 34.12 | 1.00 | 0.00 | *** |
| ot3:RepetitionPairCount | 0.22 | 1.00 | 0.64 | |
| Age:RepetitionPairCount | 1.05 | 1.00 | 0.30 | |
| ot1:Age:RepetitionPairCount | 3.33 | 1.00 | 0.07 | . |
| ot2:Age:RepetitionPairCount | 24.64 | 1.00 | 0.00 | *** |
| ot3:Age:RepetitionPairCount | 0.39 | 1.00 | 0.53 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months and Repetition Pair Count. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.8:** Regression results for Comprehension trials at 18 months during Familiarisation

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 133.49 | 1.00 | 0.00 | *** |
| ot1 | 0.23 | 1.00 | 0.63 | |
| ot2 | 2.78 | 1.00 | 0.10 | . |
| ot3 | 0.07 | 1.00 | 0.79 | |
| ot4 | 0.12 | 1.00 | 0.73 | |
| Age | 0.93 | 1.00 | 0.34 | |
| ot1:Age | 0.08 | 1.00 | 0.78 | |
| ot2:Age | 0.11 | 1.00 | 0.74 | |
| ot3:Age | 0.20 | 1.00 | 0.65 | |
| ot4:Age | 0.00 | 1.00 | 0.96 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic), ot3 (cubic) and quartic (ot4) and Age in months. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.9:** Regression results for Comprehension trials at 18 months during Test

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 5.84 | 1.00 | 0.02 | * |
| ot1 | 4.21 | 1.00 | 0.04 | * |
| ot2 | 0.37 | 1.00 | 0.54 | |
| ot3 | 0.00 | 1.00 | 0.96 | |
| ot4 | 0.01 | 1.00 | 0.93 | |
| Age | 10.13 | 1.00 | 0.00 | ** |
| ot1:Age | 7.48 | 1.00 | 0.01 | ** |
| ot2:Age | 0.20 | 1.00 | 0.66 | |
| ot3:Age | 0.12 | 1.00 | 0.73 | |
| ot4:Age | 0.06 | 1.00 | 0.81 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic), ot3 (cubic) and quartic (ot4) and Age in months. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.10:** Regression results for Nouns on Comprehension Trials split by difficulty at 18 months during Test

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 0.09 | 1.00 | 0.77 | |
| ot1 | 0.94 | 1.00 | 0.33 | |
| ot2 | 0.11 | 1.00 | 0.74 | |
| ot3 | 0.21 | 1.00 | 0.64 | |
| Age | 0.16 | 1.00 | 0.69 | |
| WordDiff | 588.16 | 2.00 | 0.00 | *** |
| ot1:Age | 1.87 | 1.00 | 0.17 | |
| ot2:Age | 0.03 | 1.00 | 0.87 | |
| ot3:Age | 0.61 | 1.00 | 0.43 | |
| ot1:WordDiff | 180.72 | 2.00 | 0.00 | *** |
| ot2:WordDiff | 3.94 | 2.00 | 0.14 | |
| ot3:WordDiff | 5.13 | 2.00 | 0.08 | . |
| Age:WordDiff | 886.95 | 2.00 | 0.00 | *** |
| ot1:Age:WordDiff | 291.79 | 2.00 | 0.00 | *** |
| ot2:Age:WordDiff | 6.80 | 2.00 | 0.03 | * |
| ot3:Age:WordDiff | 7.46 | 2.00 | 0.02 | * |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months and Word Difficulty (easy, moderate and difficult). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.11:** Regression results for Verbs and Adjectives on Comprehension Trials at 18 months during Test

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 2.45 | 1.00 | 0.12 | |
| ot1 | 4.82 | 1.00 | 0.03 | * |
| ot2 | 0.17 | 1.00 | 0.68 | |
| ot3 | 0.03 | 1.00 | 0.86 | |
| Age | 2.94 | 1.00 | 0.09 | . |
| WordType | 2.34 | 1.00 | 0.13 | |
| ot1:Age | 8.06 | 1.00 | 0.01 | ** |
| ot2:Age | 0.40 | 1.00 | 0.53 | |
| ot3:Age | 0.21 | 1.00 | 0.65 | |
| ot1:WordType | 70.57 | 1.00 | 0.00 | *** |
| ot2:WordType | 1.06 | 1.00 | 0.30 | |
| ot3:WordType | 21.03 | 1.00 | 0.00 | *** |
| Age:WordType | 5.51 | 1.00 | 0.02 | * |
| ot1:Age:WordType | 100.98 | 1.00 | 0.00 | *** |
| ot2:Age:WordType | 10.16 | 1.00 | 0.00 | ** |
| ot3:Age:WordType | 23.22 | 1.00 | 0.00 | *** |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months and Word Type (verbs and adjectives) including all difficulties. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.12:** Regression results for Novelty Bias at 18 months during Referent Selection Familiarisation

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 17.46 | 1.00 | 0.00 | *** |
| ot1 | 0.26 | 1.00 | 0.61 | |
| ot2 | 0.17 | 1.00 | 0.68 | |
| ot3 | 0.00 | 1.00 | 0.96 | |
| Age | 0.10 | 1.00 | 0.75 | |
| FamiliarityImage | 77.47 | 1.00 | 0.00 | *** |
| ot1:Age | 0.21 | 1.00 | 0.65 | |
| ot2:Age | 0.11 | 1.00 | 0.74 | |
| ot3:Age | 0.02 | 1.00 | 0.88 | |
| ot1:FamiliarityImage | 2.01 | 1.00 | 0.16 | |
| ot2:FamiliarityImage | 0.28 | 1.00 | 0.59 | |
| ot3:FamiliarityImage | 0.69 | 1.00 | 0.41 | |
| Age:FamiliarityImage | 61.11 | 1.00 | 0.00 | *** |
| ot1:Age:FamiliarityImage | 2.11 | 1.00 | 0.15 | |
| ot2:Age:FamiliarityImage | 0.52 | 1.00 | 0.47 | |
| ot3:Age:FamiliarityImage | 0.89 | 1.00 | 0.35 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months and Familiarity Image (familiar images that appeared in Reinforcement versus Comprehension trials). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.13:** Regression results for Referent Selection at 18 months during Test

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 0.36 | 1.00 | 0.55 | |
| ot1 | 0.78 | 1.00 | 0.38 | |
| ot2 | 4.49 | 1.00 | 0.03 | * |
| ot3 | 0.45 | 1.00 | 0.50 | |
| Age | 7.79 | 1.00 | 0.01 | ** |
| WordType | 2.91 | 1.00 | 0.09 | . |
| ot1:Age | 1.92 | 1.00 | 0.17 | |
| ot2:Age | 5.20 | 1.00 | 0.02 | * |
| ot3:Age | 1.84 | 1.00 | 0.17 | |
| ot1:WordType | 1.36 | 1.00 | 0.24 | |
| ot2:WordType | 5.41 | 1.00 | 0.02 | * |
| ot3:WordType | 1.59 | 1.00 | 0.21 | |
| Age:WordType | 3.50 | 1.00 | 0.06 | . |
| ot1:Age:WordType | 1.67 | 1.00 | 0.20 | |
| ot2:Age:WordType | 5.29 | 1.00 | 0.02 | * |
| ot3:Age:WordType | 1.67 | 1.00 | 0.20 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months and Word Type target (novel versus familiar). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.14:** Regression results for Retention at 18 months during Familiarisation

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 4.52 | 1.00 | 0.03 | * |
| ot1 | 3.81 | 1.00 | 0.05 | . |
| ot2 | 0.60 | 1.00 | 0.44 | |
| ot3 | 0.74 | 1.00 | 0.39 | |
| Age | 3.47 | 1.00 | 0.06 | . |
| ot1:Age | 4.56 | 1.00 | 0.03 | * |
| ot2:Age | 0.81 | 1.00 | 0.37 | |
| ot3:Age | 0.84 | 1.00 | 0.36 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic) and Age in months. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.15:** Regression results for Retention at 18 months during Test

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 2.77 | 1.00 | 0.10 | . |
| ot1 | 0.09 | 1.00 | 0.76 | |
| ot2 | 1.15 | 1.00 | 0.28 | |
| ot3 | 3.57 | 1.00 | 0.06 | . |
| Age | 3.58 | 1.00 | 0.06 | . |
| ot1:Age | 0.12 | 1.00 | 0.72 | |
| ot2:Age | 1.68 | 1.00 | 0.19 | |
| ot3:Age | 4.32 | 1.00 | 0.04 | * |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic) and Age in months. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.16:** Regression results for ELP Retest at 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 70.46 | 1.00 | 0.00 | *** |
| ot1 | 183.64 | 1.00 | 0.00 | *** |
| ot2 | 53.36 | 1.00 | 0.00 | *** |
| ot3 | 0.65 | 1.00 | 0.42 | |
| Age | 41.53 | 1.00 | 0.00 | *** |
| TrialType | 555.98 | 3.00 | 0.00 | *** |
| ot1:Age | 182.81 | 1.00 | 0.00 | *** |
| ot2:Age | 55.51 | 1.00 | 0.00 | *** |
| ot3:Age | 0.28 | 1.00 | 0.59 | |
| ot1:TrialType | 1176.15 | 3.00 | 0.00 | *** |
| ot2:TrialType | 163.04 | 3.00 | 0.00 | *** |
| ot3:TrialType | 0.74 | 3.00 | 0.86 | |
| Age:TrialType | 525.75 | 3.00 | 0.00 | *** |
| ot1:Age:TrialType | 1076.09 | 3.00 | 0.00 | *** |
| ot2:Age:TrialType | 162.82 | 3.00 | 0.00 | *** |
| ot3:Age:TrialType | 0.56 | 3.00 | 0.91 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months and Test Type. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.17:** Regression results for ELP Reinforcement Test-Retest at 18 and 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 0.02 | 1.00 | 0.90 | |
| ot1 | 0.00 | 1.00 | 0.98 | |
| ot2 | 2.52 | 1.00 | 0.11 | |
| ot3 | 0.20 | 1.00 | 0.66 | |
| Age | 0.20 | 1.00 | 0.66 | |
| RepetitionPairCount | 0.07 | 1.00 | 0.79 | |
| TestType | 0.92 | 1.00 | 0.34 | |
| ot1:Age | 0.58 | 1.00 | 0.44 | |
| ot2:Age | 2.19 | 1.00 | 0.14 | |
| ot3:Age | 0.46 | 1.00 | 0.50 | |
| ot1:RepetitionPairCount | 2.59 | 1.00 | 0.11 | |
| ot2:RepetitionPairCount | 97.30 | 1.00 | 0.00 | *** |
| ot3:RepetitionPairCount | 2.58 | 1.00 | 0.11 | |
| Age:RepetitionPairCount | 0.04 | 1.00 | 0.84 | |
| ot1:TestType | 0.25 | 1.00 | 0.62 | |
| ot2:TestType | 0.84 | 1.00 | 0.36 | |
| ot3:TestType | 0.76 | 1.00 | 0.38 | |
| Age:TestType | 0.89 | 1.00 | 0.34 | |
| RepetitionPairCount:TestType | 0.37 | 1.00 | 0.54 | |
| ot1:Age:RepetitionPairCount | 0.43 | 1.00 | 0.51 | |
| ot2:Age:RepetitionPairCount | 64.19 | 1.00 | 0.00 | *** |
| ot3:Age:RepetitionPairCount | 0.45 | 1.00 | 0.50 | |
| ot1:Age:TestType | 0.56 | 1.00 | 0.46 | |
| ot2:Age:TestType | 1.17 | 1.00 | 0.28 | |
| ot3:Age:TestType | 1.06 | 1.00 | 0.30 | |
| ot1:RepetitionPairCount:TestType | 88.48 | 1.00 | 0.00 | *** |
| ot2:RepetitionPairCount:TestType | 96.99 | 1.00 | 0.00 | *** |
| ot3:RepetitionPairCount:TestType | 14.75 | 1.00 | 0.00 | *** |
| Age:RepetitionPairCount:TestType | 0.30 | 1.00 | 0.58 | |
| ot1:Age:RepetitionPairCount:TestType | 97.53 | 1.00 | 0.00 | *** |
| ot2:Age:RepetitionPairCount:TestType | 117.08 | 1.00 | 0.00 | *** |
| ot3:Age:RepetitionPairCount:TestType | 14.32 | 1.00 | 0.00 | *** |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months, Repetition Pair Count and Test Type. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.18:** Regression results for ELP Comprehension Test-Retest at 18 and 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 9.87 | 1.00 | 0.00 | ** |
| ot1 | 6.32 | 1.00 | 0.01 | * |
| ot2 | 0.04 | 1.00 | 0.84 | |
| ot3 | 7.43 | 1.00 | 0.01 | ** |
| ot4 | 0.64 | 1.00 | 0.42 | |
| TestType | 1.03 | 1.00 | 0.31 | |
| Age | 15.69 | 1.00 | 0.00 | *** |
| ot1:TestType | 0.54 | 1.00 | 0.46 | |
| ot2:TestType | 0.20 | 1.00 | 0.65 | |
| ot3:TestType | 1.84 | 1.00 | 0.17 | |
| ot4:TestType | 0.23 | 1.00 | 0.63 | |
| ot1:Age | 9.90 | 1.00 | 0.00 | ** |
| ot2:Age | 0.02 | 1.00 | 0.89 | |
| ot3:Age | 9.64 | 1.00 | 0.00 | ** |
| ot4:Age | 0.61 | 1.00 | 0.43 | |
| TestType:Age | 1.86 | 1.00 | 0.17 | |
| ot1:TestType:Age | 1.05 | 1.00 | 0.31 | |
| ot2:TestType:Age | 0.28 | 1.00 | 0.60 | |
| ot3:TestType:Age | 2.93 | 1.00 | 0.09 | . |
| ot4:TestType:Age | 0.38 | 1.00 | 0.54 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months and Test Type. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.19:** Regression results for Nouns on ELP Comprehension Test-Retest at 18 and 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 4.56 | 1.00 | 0.03 | * |
| ot1 | 0.68 | 1.00 | 0.41 | |
| ot2 | 0.84 | 1.00 | 0.36 | |
| ot3 | 2.96 | 1.00 | 0.09 | . |
| ot4 | 0.04 | 1.00 | 0.84 | |
| TestType | 0.24 | 1.00 | 0.62 | |
| Age | 6.20 | 1.00 | 0.01 | * |
| WordDiff | 185.05 | 2.00 | 0.00 | *** |
| ot1:TestType | 0.87 | 1.00 | 0.35 | |
| ot2:TestType | 3.57 | 1.00 | 0.06 | . |
| ot3:TestType | 2.80 | 1.00 | 0.09 | . |
| ot4:TestType | 5.85 | 1.00 | 0.02 | * |
| ot1:Age | 1.25 | 1.00 | 0.26 | |
| ot2:Age | 0.43 | 1.00 | 0.51 | |
| ot3:Age | 3.66 | 1.00 | 0.06 | . |
| ot4:Age | 0.38 | 1.00 | 0.54 | |
| TestType:Age | 0.50 | 1.00 | 0.48 | |
| ot1:WordDiff | 55.75 | 2.00 | 0.00 | *** |
| ot2:WordDiff | 163.32 | 2.00 | 0.00 | *** |
| ot3:WordDiff | 3.52 | 2.00 | 0.17 | |
| ot4:WordDiff | 67.63 | 2.00 | 0.00 | *** |
| TestType:WordDiff | 412.97 | 2.00 | 0.00 | *** |
| Age:WordDiff | 285.42 | 2.00 | 0.00 | *** |
| ot1:TestType:Age | 0.65 | 1.00 | 0.42 | |
| ot2:TestType:Age | 3.01 | 1.00 | 0.08 | . |
| ot3:TestType:Age | 3.26 | 1.00 | 0.07 | . |
| ot4:TestType:Age | 3.74 | 1.00 | 0.05 | . |
| ot1:TestType:WordDiff | 175.97 | 2.00 | 0.00 | *** |
| ot2:TestType:WordDiff | 89.17 | 2.00 | 0.00 | *** |
| ot3:TestType:WordDiff | 155.97 | 2.00 | 0.00 | *** |
| ot4:TestType:WordDiff | 9.31 | 2.00 | 0.01 | ** |
| ot1:Age:WordDiff | 107.93 | 2.00 | 0.00 | *** |
| ot2:Age:WordDiff | 181.98 | 2.00 | 0.00 | *** |
| ot3:Age:WordDiff | 2.20 | 2.00 | 0.33 | |
| ot4:Age:WordDiff | 88.81 | 2.00 | 0.00 | *** |
| TestType:Age:WordDiff | 467.72 | 2.00 | 0.00 | *** |
| ot1:TestType:Age:WordDiff | 201.94 | 2.00 | 0.00 | *** |
| ot2:TestType:Age:WordDiff | 93.18 | 2.00 | 0.00 | *** |
| ot3:TestType:Age:WordDiff | 139.24 | 2.00 | 0.00 | *** |
| ot4:TestType:Age:WordDiff | 1.97 | 2.00 | 0.37 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic), ot3 (cubic) and quartic (ot4), Age in months, Word Difficulty and Test Type. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.20:** Regression results for Verbs and Adjectives on ELP Comprehension Test-Retest at 18 and 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 0.00 | 1.00 | 0.96 | |
| ot1 | 6.13 | 1.00 | 0.01 | * |
| ot2 | 0.70 | 1.00 | 0.40 | |
| ot3 | 2.03 | 1.00 | 0.15 | |
| ot4 | 1.73 | 1.00 | 0.19 | |
| Age | 0.10 | 1.00 | 0.76 | |
| WordType | 190.54 | 1.00 | 0.00 | *** |
| TestType | 0.48 | 1.00 | 0.49 | |
| ot1:Age | 8.18 | 1.00 | 0.00 | ** |
| ot2:Age | 1.13 | 1.00 | 0.29 | |
| ot3:Age | 2.82 | 1.00 | 0.09 | . |
| ot4:Age | 2.17 | 1.00 | 0.14 | |
| ot1:WordType | 42.67 | 1.00 | 0.00 | *** |
| ot2:WordType | 0.83 | 1.00 | 0.36 | |
| ot3:WordType | 18.88 | 1.00 | 0.00 | *** |
| ot4:WordType | 12.29 | 1.00 | 0.00 | *** |
| Age:WordType | 290.83 | 1.00 | 0.00 | *** |
| ot1:TestType | 0.00 | 1.00 | 0.98 | |
| ot2:TestType | 1.59 | 1.00 | 0.21 | |
| ot3:TestType | 1.00 | 1.00 | 0.32 | |
| ot4:TestType | 0.22 | 1.00 | 0.64 | |
| Age:TestType | 0.69 | 1.00 | 0.41 | |
| WordType:TestType | 2.33 | 1.00 | 0.13 | |
| ot1:Age:WordType | 49.40 | 1.00 | 0.00 | *** |
| ot2:Age:WordType | 1.30 | 1.00 | 0.25 | |
| ot3:Age:WordType | 20.38 | 1.00 | 0.00 | *** |
| ot4:Age:WordType | 8.65 | 1.00 | 0.00 | ** |
| ot1:Age:TestType | 0.07 | 1.00 | 0.80 | |
| ot2:Age:TestType | 1.72 | 1.00 | 0.19 | |
| ot3:Age:TestType | 1.13 | 1.00 | 0.29 | |
| ot4:Age:TestType | 0.48 | 1.00 | 0.49 | |
| ot1:WordType:TestType | 64.49 | 1.00 | 0.00 | *** |
| ot2:WordType:TestType | 26.13 | 1.00 | 0.00 | *** |
| ot3:WordType:TestType | 0.92 | 1.00 | 0.34 | |
| ot4:WordType:TestType | 28.43 | 1.00 | 0.00 | *** |
| Age:WordType:TestType | 29.52 | 1.00 | 0.00 | *** |
| ot1:Age:WordType:TestType | 46.13 | 1.00 | 0.00 | *** |
| ot2:Age:WordType:TestType | 29.41 | 1.00 | 0.00 | *** |
| ot3:Age:WordType:TestType | 4.57 | 1.00 | 0.03 | * |
| ot4:Age:WordType:TestType | 27.11 | 1.00 | 0.00 | *** |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic), ot3 (cubic) and quartic (ot4), Age in months, Word Type and Test Type. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 2.21:** Regression results for ELP Retention Test-Retest at 18 and 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 0.11 | 1.00 | 0.74 | |
| ot1 | 1.18 | 1.00 | 0.28 | |
| ot2 | 0.02 | 1.00 | 0.88 | |
| ot3 | 5.90 | 1.00 | 0.01 | * |
| TestType | 0.52 | 1.00 | 0.47 | |
| Age | 0.17 | 1.00 | 0.68 | |
| ot1:TestType | 0.02 | 1.00 | 0.90 | |
| ot2:TestType | 0.00 | 1.00 | 0.99 | |
| ot3:TestType | 0.67 | 1.00 | 0.41 | |
| ot1:Age | 1.45 | 1.00 | 0.23 | |
| ot2:Age | 0.04 | 1.00 | 0.83 | |
| ot3:Age | 6.74 | 1.00 | 0.01 | ** |
| TestType:Age | 0.48 | 1.00 | 0.49 | |
| ot1:TestType:Age | 0.00 | 1.00 | 0.96 | |
| ot2:TestType:Age | 0.00 | 1.00 | 0.99 | |
| ot3:TestType:Age | 1.67 | 1.00 | 0.20 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months and Test Type. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

# Chapter 3

# Early Language Input in Relation to Language Outcomes

## 3.1 Introduction

Children show large individual variability in vocabulary size and rate of development at early stages of language learning (Fenson et al., 1994). Individual differences early in development relate to later language abilities (e.g., Cristia et al., 2014), with vocabulary size during the second year of life being one of the best predictors of success in school (e.g., Durham, Farkas, Hammer, Bruce Tomblin, & Catts, 2007; Farkas & Beron, 2004). It is important to understand what variables have an impact on individual differences in early vocabulary. A possible explanation for individual variability in children's vocabulary and language development, is children's different experiences with language. In fact, research suggests that environmental variables play a prominent role (see Hoff, 2006 for a review). Ensuring all children reach their full potential requires understanding what factors impact language development and the mechanisms that support early word learning. A better understanding of what factors are important for language acquisition and how they relate to different language processing mechanisms, will support efforts to boost children's language development abilities when needed.

Multiple studies have highlighted the role of language input as a key variable that has an impact on children's language development (see Rowe & Weisleder, 2020 for a review). There is a large body of research investigating how both quantity (e.g., number of words) and quality (e.g., lexical diversity, sentence complexity) of language input are related to later language abilities and literacy skills (e.g., Huttenlocher et al., 1991; Hirsh-Pasek et al., 2015; Rodriguez & Tamis-Lemonda, 2011; Rowe, 2012). Moreover, language input has been related to other non-linguistic cognitive capacities such as executive functioning (Sarsour et al., 2011), and social skills (Connell & Prinz, 2002).

It is less well documented how early language input relates to the basic cognitive processes that support early word learning. It is possible that this is due to the fact that language input is typically related to parental checklists rather than to direct measures of word learning. Parental checklists give measures of total vocabulary but they do not measure the language processes involved in the word learning task. Moreover, most of the extant literature focuses on data from a single time point. When measures of language input and language outcomes are gathered at the same moment in time, it is hard to know which variable is the predictor and which one is the predicted. This can be told apart using longitudinal studies, or studies gathering input measures and language processing measures at different time points, however, those are relatively scarce.

The goal of this study was to relate children's early language experiences to their basic cognitive processes that support early word learning. In particular, we investigated the relationship between home language input and the underlying cognitive and language processes that support language development. Those processes include word processing speed, comprehension abilities, novelty biases, referent selection and retention of new words. To understand how language input influences developing language processes at different time points, we used longitudinal data relating early children's language environment to later linguistic processing abilities.

### 3.1.1 Language Input and Language Outcomes

Studies have shown that properties of linguistic input to children – its quantity and quality – are related to early language development (e.g., Huttenlocher et al., 1991, 2010; Weisleder & Fernald, 2013; Hurtado et al., 2008a). These effects have been found across a variety of tasks such as in-lab experiments, play sessions and naturalistic home recordings. Studies show that these effects also change over time, and that children benefit from different aspects of quantity and quality of caregiver input at different points in development (Rowe, 2012).

Exposure to language input also differs between social classes (or socioeconomic status, SES) in Western contexts. These differences have been shown to explain some of the socioeconomic differences in young children's vocabulary skills (e.g., Hoff, 2003). In a study with middle-class families, Huttenlocher et al. (1991) found that only the amount of parental input predicted vocabulary growth between 14 and 26 months. In their seminal study, Hart and Risley (1995), looked at variation in the quantity of input across the early childhood period in families ranging in SES. They estimated that by the time children reach school age, those growing up in higher-SES families were, on average, exposed to 30 million more words than children growing up in lower-SES families. However, there is some recent counter evidence of these findings proposing that parental linguistic input may be a limited indicator for certain groups such as lower SES families (Sperry, Sperry, & Miller, 2019; Sperry et al., 2019), although there is some debate around this last argument (Golinkoff, Hoff, Rowe, Tamis-LeMonda, & Hirsh-Pasek, 2019). Hart and Risley (1992, 1995) also found strong positive associations between quantity of caregiver input and children's vocabulary growth, supporting the notion that the quantity of parental vocabulary input influences children's rate of vocabulary growth. Quality of input also mattered with higher-SES parents responding more to their children, producing more affirmative and encouraging instances and fewer prohibitions. In addition, high SES parents showed more diverse input because they produced more noun types and modifiers per hour. Similar findings were reported in a Family Life study with a large homogeneous sample of 1,292 children followed from birth (Vernon-Feagans et al., 2020). The authors found that maternal language input indexed as the number of different words, mean length of utterance and number of wh-questions, partially mediated the relationship between maternal education and later child language at school age. These studies show evidence that language input, particularly quantity, is associated with later language abilities.

Multiple studies also report the qualitative properties of language input also influence language development (see Rowe & Snow, 2020 for a review). There are many features that can contribute to rich language input. Examples include the diversity of words children hear in their input (Hurtado et al., 2008b; Huttenlocher et al., 2010; Weisleder & Fernald, 2013), the use of sophisticated vocabulary (Weizman & Snow, 2001), talker variability (Rost & McMurray, 2009), the number of words produced in isolation (Brent & Siskind, 2001), parental responsiveness in interactions (Tomasello & Farrar, 1986), or referential transparency in the words that children hear (Cartmill et al., 2013; Bergelson & Aslin, 2017). It is likely that both quantity and quality of language input help children build their language abilities.

In fact, some studies examining both quantity and quality of language input have shown that they both contribute to later language skill but in different ways. A study by Hoff and Naigles (2002) showed that 2-year-old children's lexical development benefited from higher quantity, more lexical richness, and more syntactic complexity of maternal language input in data from mother-child conversation during dyadic play. Benefits from both quantity and quality of input have also been found across SES. Huttenlocher et al. (2010) followed a group of children with diverse SES backgrounds longitudinally from 14 to 46 months to examine the role of quantity of input (e.g., word tokens) and diversity of input (i.e., variety of words and syntactic structures) in children's vocabulary and syntactic growth. In this study, variations in language input, particularly differences in the syntactic structures caregivers used, affected children's language growth. Further, while quantity and diversity of input was related to SES, diversity of caregiver speech was a significant predictor of child vocabulary growth, measured as the word types children produced when controlling for SES.

Even though both quantity and quality of language input are beneficial for children's language skills, they contribute to children's learning abilities in different ways across development. Rowe (2012) examined quantitative and qualitative proprieties of caregiver input in a longitudinal sample of parent–child dyads to determine which aspects of input contribute most to children's vocabulary skill across early development. Input measures from parent–child interactions at 18, 30, and 42 months were examined in relation to children's vocabulary skill one year later, when children were 30, 42, and 54 months. Input quantity and children's previous vocabulary skill explained variation in later vocabulary ability when SES was controlled. However, specific measures of input quality related to child vocabulary skill at different points in development, even when controlling for SES and input quantity. This means that, even though quantity of input was predictive of later vocabulary, more fine-grained aspects of input (i.e., qualitative aspects) did matter for language skill and were dependent on the child's age or language ability. These results suggest that quantity and quality of input contribute in different ways to language skills across development because quantity of input was most important for vocabulary skills during the 2nd year of life, quality of input such as the diversity or sophistication of the vocabulary in the input, was most important during the 3rd year of life, and the use of decontextualized language such as narrative and explanations in the input was most beneficial during the 4th year of life.

In fact, several studies suggest that certain linguistic features of the input might be more or less helpful at different points of children's development (see Rowe & Snow, 2020 for a review). For example, in infancy, words in isolation are more easily learned (Brent & Siskind, 2001) even when controlling for frequency of occurrence (Swingley & Humphrey, 2018). Variation in utterance length (i.e., hearing both short and long utterances) is also beneficial during infancy. However, later on in toddlerhood, this pattern reverses because children process words that occur in sentence frames faster than those that occur in isolation (A. Fernald & Hurtado, 2006). Once children reach preschool, studies show that complex syntax boosts children's syntactic development (Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002). Thus, it is possible that quantity of language input is more relevant early in development (i.e., from infancy to toddlerhood), helping to initiate the language learning process, whereas quality of input, in different forms, might be more relevant later in language development (i.e., from toddlerhood to preschool ages), when children are able to benefit from the richness of that speech (Golinkoff et al., 2019; Rowe, 2012; Rowe & Snow, 2020). It is worth noting however, that both quantity and quality are not independent constructs because parents that use larger quantities of speech also have more chances to use more diverse language (Hoff & Naigles, 2002). This makes it is difficult to isolate the influence of each variable on language acquisition. To overcome this problem, computational work controlled the input to a computational model so that quantity of exposure to linguistic input and the quality of that input (lexical diversity) were independently manipulated (Jones & Rowland, 2017). On this work, the model was tested on input that was artificially manipulated to increase quantity (keeping lexical diversity constant) as well as on input that was artificially manipulated to increase lexical diversity (keeping quantity constant). The model trained on input quantity consistently showed an initial advantage early in learning. However, the model trained on input with higher lexical diversity quickly superseded the model trained on quantity of input, providing a superior learning environment by the end of the learning process, a prediction that was also confirmed against children's data. The model trained on a lexically diverse input also performed better on non-word repetition, sentence recall tests and was quicker at learning new words (Jones & Rowland, 2017). This shows that while input quantity may be important early in learning, lexical diversity is ultimately more crucial.

A further question is *how* input influences later vocabulary skills. One hypothesis is that input might influence vocabulary growth via the child's language processing abilities. Thus, language input is beneficial for children who are able to efficiently process that input. That might create a cascade effect were early input boosts language abilities to which parents respond with more sophisticated input. Hurtado et al. (2008a) examined maternal speech at 18 months in relation to children's speech processing efficiency and vocabulary at 18 and 24 months. Children whose mothers provided larger quantities of language input at 18 months knew more words and were faster in word recognition at 24 months. However, the influences of caregiver speech on speed of word recognition and vocabulary were largely overlapping. Thus, it is not clear in this study if the relation between language experience and processing efficiency could be explained by children's vocabulary size. It is possible that young children who are exposed to larger quantities of linguistic input have more opportunities to develop and build language processing skills which could facilitate language development. To further explore this possibility, a follow up study with a sample of low-income Spanish-speaking US families, measured the link between early language experience and language processing efficiency, while exploring whether processing skills mediate the relation between early language experience and later vocabulary knowledge. In this work, Weisleder and Fernald (2013) found that 19-month-old infants who experienced more child-directed speech were also more efficient at processing familiar spoken words. Moreover, those children had larger expressive vocabularies at 24 months. Importantly, in this study speech overheard by the child was unrelated to vocabulary outcomes, and children's lexical processing abilities mediated the effect of child-directed input on future expressive vocabulary. This led the authors to conclude that larger quantities of child-directed input provided children more opportunities to practice recognizing words, which led to greater processing efficiency, facilitating word learning. Interestingly, even within this quite homogeneous sample of lower-SES families, large variation in parent input was found and was predictive of child vocabulary growth. Similar evidence has also been found in older children.

Mahr and Edwards (2018), used looking responses in preschool age children to measure how lexical processing as well as language input at 28–39 months, predicted vocabulary size (assessed using direct measures) one year later. The authors found that language input and lexical processing predicted receptive vocabulary growth, indicating that both language experience as well as processing abilities are related to vocabulary development.

These studies provide examples of how language input relates to later language ability as a function of the child's linguistic and cognitive capacities, particularly speed of word processing. There is evidence that input is also associated with children's segmentation abilities. A cognitive skill children need to have to use language input to build vocabulary is the ability to extract words from running speech. Newman, Rowe, and Bernstein Ratner (2016) related quality of language input, together with children's ability to segment words from speech in infancy to later vocabulary outcomes. Input quality was indexed by the lexical properties of maternal child-directed speech to 7-month-old infants. Infants' abilities to segment lexical targets from conversational child-directed utterances was measured in an experimental paradigm, as well as vocabulary outcomes at age 24 months, measured via parental checklist. Both repetitiveness (i.e., type–token ratio) in maternal input and the child's speech segmentation skills at 7 months independently predicted language outcomes at 24 months. This literature shows the role of language input on the child's linguistic and cognitive capacities. However, not many studies have documented these relationships and none looks at how input affects multiple cognitive abilities related to word learning in the same children.

### 3.1.2 Conversational Experience as a Measure of Language Input

Language acquisition occurs within a social context (Tomasello, 2019). Infants engage in communication using eye gaze, gestures, and vocalisations before they utter their first words (e.g., Snow, 1977). This exchange, called the *conversational duet*, has been proposed to play a key role for language and socio-cognitive development (e.g., Song, Spier, & Tamis-Lemonda, 2014). In fact, some literature measuring conversational turns has suggested that this metric has a stronger relationship to children's language outcomes than the mere quantity of language input (e.g., Gilkerson et al., 2018; Zimmerman et al., 2009).

Studies using Language Environment Analysis (LENA) technology (Ford, Baer, Xu, Yapanel, & Gray, 2008) have linked both quantity of adult input and conversational experience to language outcomes with data showing that children's conversational turn count (CTC), and to a lesser extent the adult word count (AWC), are associated with language outcomes. The LENA technology is a composite recording and analysis package that records up to 16 hours of a child's language environment across one day (see Greenwood, Bourque, & Buzhardt, 2011 for an extension of Hart and Risley (1995) study using this technology). The LENA software provides estimates of the number of adult words a child hears (AWC). It also computes the number of conversational turns (CTC) in which a child engages and the number of child vocalizations (CVC). In addition, LENA provides measures such as the amount of extraneous background noise a child hears due to electronic media such as television and radio. Although the LENA system was originally developed in American English, the device has been used in previous literature showing high and moderate estimates in several other languages such as Spanish (Weisleder &

Fernald, 2013), Swedish (Schwarz et al., 2017), Korean (McDonald, Kwon, Kim, Lee, & Ko, 2021), Chinese (Gilkerson et al., 2015), European French (Canault, Le Normand, Foudil, Loundon, & Thai-Van, 2016) and even bilingual contexts (Orena, Byers-Heinlein, & Polka, 2019). Although these studies report acceptable rates of overlap between the LENA output and manual annotation of the same recordings in those languages, some recent studies report lower accuracy for CTC in comparison to AWC and CVC (Cristia et al., 2021).

Using the LENA system, Donnelly and Kidd (2021) measured conversational turns in relation to vocabulary outcomes in a longitudinal sample of children followed from 9 to 24 months of age. Day-long home audio recordings provided the number of conversational turns. Vocabulary was measured independently via parental report. Growth curve analyses revealed a bidirectional relationship between conversational turns and vocabulary growth, controlling for the amount of words in children's environments, suggesting that social interaction in the form of conversational turns is an important component of early language acquisition.

In a cross-sectional study of children aged 2–48 months, Gilkerson et al. (2017) found that both AWC and CTC were significantly correlated with several outcome measures of language and cognitive development. In a follow up study, the authors tested a subset of the children on language outcome measures 10 years later, finding that only conversational turn measures early in development related to language outcomes (Gilkerson et al., 2018). Similarly, a study investigating the relationship between CTC (measured by LENA) and language development in 2–36-month-olds found that conversational turns predicted language proficiency, but not vice versa (Zimmerman et al., 2009).

Romeo et al. (2018) showed that the number of conversational turns in daylong LENA recordings predicted 4- to 6-year-old children's language proficiency over and above measures of input quantity and was associated with greater activation in Broca's Area (the Left Inferior Frontal Gyrus) during a language processing task conducted during an MRI scan. In this study, turn counts, and not adult input, were also associated with greater myelin concentrations in language related areas of the brain such as the left arcuate fasciculus, and the superior longitudinal fasciculus, independently of SES status. Finally, in a recent meta-analysis of effect sizes across 13 studies, Wang et al. (2020) reported a moderate effect size (r = 0.32) between LENA turn measures and language proficiency with smaller effects regarding the relationship between LENA adult input and language proficiency (r = 0.21).

It is not surprising that conversational turns are relevant for language development. Research on children's language environments, primarily in Western contexts, is converging on the idea that it is not merely the quantity of child-directed speech that best predicts language learning, but also the extent to which that speech occurs in episodes of joint engagement and attention (Hirsh-Pasek et al., 2015). This joint engagement typical of conversational turns could be conceptualized as an important aspect of the quality of language input.

The reviewed literature shows evidence that early language experiences are related to children's later language abilities. Both quantity and quality of language input have been related to vocabulary and language skills, with children benefiting from different aspects of their input at different points in development. Recent technological advances and automatised techniques, such the LENA system, allow easier to quantification of different aspects of children's language input. This allows more efficient examination of the relationships between the amount of adult words and conversational turns, and children's language abilities.

However, most prior studies have investigated the relationships between language input and language abilities using indirect measures of language skill, which do not capture children's language processing abilities. Thus, it is not clear how early input relates to the language processes involved in word learning. As noted above, there is some evidence that greater amounts of input are associated with faster speed of word processing. However, the task of learning a new word relies on several cognitive capacities. When faced with a new word, novelty biases help children select the correct referent and remember the association between the new word and referent. This task will be easier if children have larger vocabularies because that will help them select the correct referent. Thus, word learning involves several skills, however, to our knowledge no study has related measures of the multiple language processing skills involved in the word learning task to language input in the same group of children. Moreover, no work has related different aspects of language input to changes in language processing abilities over development.

### 3.1.3 The Present Study

To have a richer picture of how the processes that support language development are impacted by language input, the present study investigates the relationship between home language input at two different time points, infancy and toddlerhood, and early language processes in a sample of UK infants. Adult input and conversational turns were measured in children's natural environment using the LENA system. Subsequently, children's language processing abilities including speed of word processing, comprehension abilities, novelty biases, referent selection and retention of new words were measured using the Early Language Processing task.

We gathered home language data at two longitudinal time points –infancy and toddlerhood. Our daylong recordings lasted up to 16 hours. Day-long recordings may yield more representative data on children's language experiences than shorter samples of language input (e.g., gathered in the lab during parent-child interactions), because they are more likely to capture language in different contexts that are more representative of children's day-to-day lives. In a recent meta-analysis by Anderson et al. (2021) the length of the observation significantly moderated the association between the quality of parental linguistic input and child language, with longer observation periods leading to larger effect sizes.

To measure language processing abilities, we tested children at 18 and at 30 months-of-age using the ELP task. Chapter 2, details how we developed this portable, efficient eye-tracking based task, that is able to capture group and individual differences in multiple language measures such as speed of word processing, word comprehension, novelty biases, referent selection and retention of new words. Here, the goal is to explore how children's language input at home influences the different ELP outcome measures.

## 3.2 Methods

The goal of this study is to assess how children's language input in infancy and toddlerhood measured with LENA influenced their language processing abilities at 18 and 30 months measured with the ELP task. Towards this aim, we collected both LENA and ELP data from the same group of children at two different time points.

### 3.2.1 Participants

The main analysis of this chapter includes children from which we have two LENA observations as well as ELP data at either 18 or 30 months. Thus, the data is a sub-sample of the ELP data reported in Chapter 2.

Demographic information for the subsample of 54 children with with two LENA observations (24 female) can be seen in Table 3.1. Overall, children were 94.4% white and 5.6% mixed race, 67.8% of mothers had completed a Bachelor's degree or higher. Median family annual income was £44,200 (range £13,000-£52,000). The first LENA measurement, "LENA Infant," was collected when the children were between 4 and 13 months of age ($M$ = 6.91 months, $SD$ = 1.69 months). The second LENA measurement, "LENA Toddler," was collected when the children were between 17 and 27 months of age ($M$ = 20.25 months, $SD$ = 2.20 months). From the initial sample of 56 children with LENA recordings at both time points, 2 children were excluded because they were much older then their peers at the time of the infant recording: 1 child was 19 months old at the point of the first LENA recording, the other was 30 months old at the point of the LENA Toddler recording. Participants had normal or corrected-to-normal vision. Inclusion criteria included (1) uncomplicated birth between 37 and 42 weeks; (2) no reports of alcohol or illicit drug use during pregnancy; (3) no familial history of major psychiatric or depressive illness; (4) no preexisting neurological conditions or major head trauma. These criteria were confirmed during parental interviews at enrollment.

We collected LENA data from an additional 35 toddlers, but these data was not included in the final analysis due to children not having data for one of the two LENA sessions (7 children missed LENA Infant but had data for LENA Toddler; 28 children missed LENA Toddler but had data for LENA Infant).

The final sample of children with with LENA Infant, LENA Toddler and

ELP data at 18 months of age includes a total of 35 children (16 girls). The final sample of children with with LENA Infant, LENA Toddler and ELP data at 30 months of age includes a total of 21 children (8 girls).

This project was reviewed and approved by the UK NHS Health Research Authority Ethics committee (Protocol ID: IRAS 196063; PI: John P. Spencer and ID: 211250 PI: Larissa K. Samuelson).

### 3.2.2 Procedure

At each LENA time point (LENA Infant and LENA Toddler), participants were given a LENA audio recording device to take home. We had between 1 and 3 days of LENA data, containing between 8 and 16 hours of data per day for each participant. In total we gathered 3337 hours of recordings for LENA Infant ($M$ = 40.69 hours per participant, $sd$ = 10.46 hours) and 1591.83 hours of recording data for LENA Toddler ($M$ = 26.09 hours per participant, $sd$ = 13.80 hours).

At approximately 18 and 30 months, the same children came to the Laboratory for an ELP session. The procedure was is reported in Chapter 2. Children sat on their parent's lap or on high chair. An Eye-Link Duo (SR Research, Ontario, Canada) eye-tracker in the remote setting captured children's gaze. During the ELP session, parents also filled in the online adaptation of the OCDI vocabulary checklist. Data related to that vocabulary checklist is also reported in Chapter 2.

Before the sessions, parents signed an informed consent form. On ELP sessions, children received a small toy of their choosing and a t-shirt for participating.

| **LENA Sample Demographics**<br>Participants with two observations N = 54 (24 girls) | |
|---|---|
| **Age in Months** | |
| **LENA Infant** | |
| Mean (SD) | 6.91 (1.69) |
| Median [Min, Max] | 6.57 [4.67, 13.77] |
| **LENA Toddler** | |
| Mean (SD) | 20.25 (2.20) |
| Median [Min, Max] | 19.55 [17.03, 27.77] |
| **Ethnicity** | |
| African | 0 (0%) |
| Asian | 0 (0%) |
| Mixed | 3 (5.6%) |
| White | 51 (94.4%) |
| Not specified | 0 (0%) |
| **Mother's Education Status** | |
| Left School | 1 (1.8%) |
| GCSE/O levels equivalent | 5 (9.2%) |
| A levels or equivalent | 6 (11.1%) |
| Trade apprenticeship | 0 (0%) |
| Some university | 5 (9.2%) |
| Bachelor's Degree | 24 (44.4%) |
| Master's Degree | 7 (12.9%) |
| Doctorate or Professional Degree | 6 (11.1%) |
| Not specified | 0 (0%) |
| **Father's Education Status** | |
| Left School | 1 (1.8%) |
| GCSE/O levels equivalent | 8 (14.8%) |
| A levels or equivalent | 7 (12.9%) |
| Trade apprenticeship | 8 (14.8%) |
| Some university | 1 (1.8%) |
| Bachelor's Degree | 18 (33.3%) |
| Master's Degree | 6 (11.1%) |
| Doctorate or Professional Degree | 5 (9.2%) |
| Not specified | 0 (0%) |

**Table 3.1:** Summary of LENA sample demographics.

### 3.2.3 Data Processing and Analytical Approach

**LENA data:** The home audio recordings were exported using the LENA proprietary software. The advanced data extraction software (ADEX) from LENA provided several estimates of the child's language environment, including Adult Word Count (AWC), defined as the number of words spoken in the vicinity of the child, Child Vocalization Count (CVC), defined as the number of vocalizations (including words and non-words, such as babbling or exclamations such as *ah!*), the child's exposure to non-social electronic media (e.g., TV, radio, music), and child–adult conversational turn count (CTC), defined as two discrete utterances between child–adult pairs that contain a pause of no longer the 5 seconds. Note that CTC is a composite measure that contains AWC and CVC that happened consecutively (Ford et al., 2008). We used the Vocalisation Activity Block (One Row per Block per Recording) including Segment details as output, because it gave the highest resolution. From each extended home recording at each LENA time point we found the hour with highest AWC, CTC and CVC by first extracting counts for 1 hour bins across the entire recording (as in Romeo et al., 2018). The maximum count per each LENA measure was then selected across the different LENA days. This gave a maximum AWC, CTC and CVC per hour for each LENA observation. This processing was done using the statistical package R (R Core Team, 2017).

We had three main questions regarding the language measures extracted from LENA (AWC, CTC and CVC): Q1) Do AWC, CTC and CVC differ across the two time points? Q2) Were AWC, CTC and CVC consistent across individuals at both time points? Q3) Were LENA measures related to SES differences in our sample?

To answer Q1 we ran a Wilcoxon test (a more conservative version of a t-test), to measure if there were differences between LENA measures at both time points. To answer Q2, we ran correlation analyses to look at the possible relationships between LENA measures during infancy and LENA measures during toddlerhood. To answer Q3, we ran linear models to assess possible relationships between LENA and SES status. SES (zSES) was computed as the average z-score (standard score) of maternal education.

In this set of models, LENA Infant and zSES predicted LENA Toddler measures. All LENA measures were centered. Model fit was assessed using the *check_model* function from the R package Performance (Lüdecke, Ben-Shachar, Patil, Waggoner, & Makowski, 2021), which generates a visual check of various model assumptions such as normality of residuals, normality of random effects, linear relationship, homogeneity of variance and multicollinearity.

In all of our LENA data analyses, we included the three LENA outcome measures (AWC, CTC and CVC), as a quality check. For example, one would expect an increase in child vocalisations as the child ages or a strong correlation between child vocalisations and conversational turn counts.

**ELP data:** The eye-tracking data from ELP were pre-processed using Data Viewer (SR-Research, Ontario, Canada). Trials were segmented into periods of interest (IP) using message-based events. Areas of interest (AOI) were set to be 50% bigger than target objects to account for calibration errors and drifts in the eye tracker. Sample reports were exported and raw gaze position was processed using the statistical package R (R Core Team, 2017).

Eye-tracking data from word onset to 1800 ms after onset from the ELP test phase was processed using the eyetrackingR package (Dink & Ferguson, 2016). During data processing, trials with more than 40% of trackloss were removed from the analysis. Mean proportions of looks to the target per each ELP trial type (Reinforcement, Comprehension, Referent Selection and Retention), as well as proportions of looks to the novel item during familiarisation of Referent Selection trials (for Novelty Bias) were extracted following the same procedure as in Chapter 2.

The main question was whether language input at two time points, early in infancy and later in toddlerhood, is associated with language processes at 18 and 30 months. To examine this, we ran a set of linear models with LENA measures at both time points predicting ELP measures at 18 months or at 30 months separately. The main analysis examining LENA and ELP relationships did not look at child vocalisations because our focus was on the linguistic input (AWC and CTC) rather than on the child's productions, CVC. We also did not look at relationships between ELP measures, LENA measures and OCDI, because our prior analyses showed strong correlations between ELP Comprehension and both OCDI production and comprehension scores (see Chapter 2). Thus, we used ELP Comprehension as the child's ability to understand and recognise words (i.e., comprehension abilities).

LENA input measures were included in each model as fixed terms and entered separately because they are highly correlated due to the fact that turn count is calculated using adult words and child vocalisations. Thus, we never modeled AWC and CTC together. Both LENA input measures were scaled and centered. The predicted ELP measures were mean proportion of looking to target in ELP Reinforcement, Comprehension, Referent Selection, and Retention trails, each modeled separately. We also included ELP Novelty Bias as an additional predicted measure, which was calculated based on mean proportion of looks to the novel image in Referent Selection trials before the child heard the target word (i.e., during the familiarisation phase of the trial). Models predicting ELP Comprehension included Word Type (noun, verb and adjective) and Word Difficulty (easy, moderate and difficult) as fixed effects.

Models predicting ELP Referent Selection included Word Type, which refers to whether the target was novel or a familiar noun as predictors. Models predicting ELP Novelty Bias, included Familiar Image Type as predictor. Familiar Image Type refers to the type of familiar image that was paired with the novel one. This image could be familiar or highly familiar. Familiar images also appeared in ELP Comprehension, highly familiar images also appeared on ELP Reinforcement. Model fit to the data was assessed using the DHARMa R package (Hartig, 2021), which uses a simulation-based approach to create readily interpretable scaled (quantile) residuals for fitted (generalized) linear mixed models, in addition the Performance R package (Lüdecke et al., 2021). For each model, the effect of each parameter was assessed with an F test, in particular, we used the ANOVA function from the car R package (R Core Team, 2017), which tests whether the model terms are significant. All the reported effects and interactions are those that remained after using this method.

## 3.3 Results

The main aim of this study was to examine whether language input predicts children's language processing abilities. We first report analysis on the LENA data only, characterizing the amount of input children heard, their own turn and vocalisation productions, and changes in these variables over time. Then we examine whether the language input the children heard as infants and toddlers predicted language processing measures at 18 and 30 months.

### 3.3.1 Language Input

The first question was whether the three LENA measures of adult input (AWC), conversational turns (CTC) and child vocalisations (CVC) differed across LENA observations (Infant versus Toddler). As can be seen in Figure 3.1, paired Wilcoxon signed-rank tests showed that adult word count significantly decreased from the infant to the toddler observation (AWC Infant $M$ = 5540, $sd$ = 2205.695; AWC Toddler $M$ = 3680.4, $sd$ = 1683.103; V = 1325, p <.000) . However the max turn count (CTC Infant $M$ = 104.6, $sd$ = 34.543; CTC Toddler $M$ = 123.22, $sd$ = 63.484; V = 491.5, p = 0.04735) and child vocalizations increased (CVC Infant $M$ = 329.5, $sd$ = 127.77; CVC Toddler $M$ = 423.4, $sd$ = 156.543; V = 472.5, p = 0.020). Thus, the three measures together suggest that in toddlerhood adults are speaking less and infants contributing more both by vocalizing more and by taking turns.

To explore the consistency of LENA measures in individuals across both time points (Q2), we ran a set of correlations. As can be seen in Figure 3.2, LENA Infant AWC was positively related with LENA Toddler AWC ($R$ = 0.379, $p$ = 0.004). This indicates that the amount of adult talk to children at both time points was consistent. There was also a positive correlation between AWC and CVC and CTC at both time points. This is not surprising because CTC is defined as AWC and CVC instances that happened consecutively without a pause longer than 5 seconds.

The last question (Q3) was whether there were similar SES effects in our sample as those reported in literature (Hart & Risley, 1995). For instance, we wanted to examine whether higher SES children were exposed to more language input. We fit three linear models, one for each LENA measure (AWC, CTC and CVC), to the LENA data with LENA Infant predicting LENA Toddler as a function of the mean standardized SES across LENA observations. Results showed no SES effects, although there was a marginal interaction between turns in the LENA Infant recordings and zSES. Our sample is quite homogeneous and since SES was not a significant predictor, we do not include

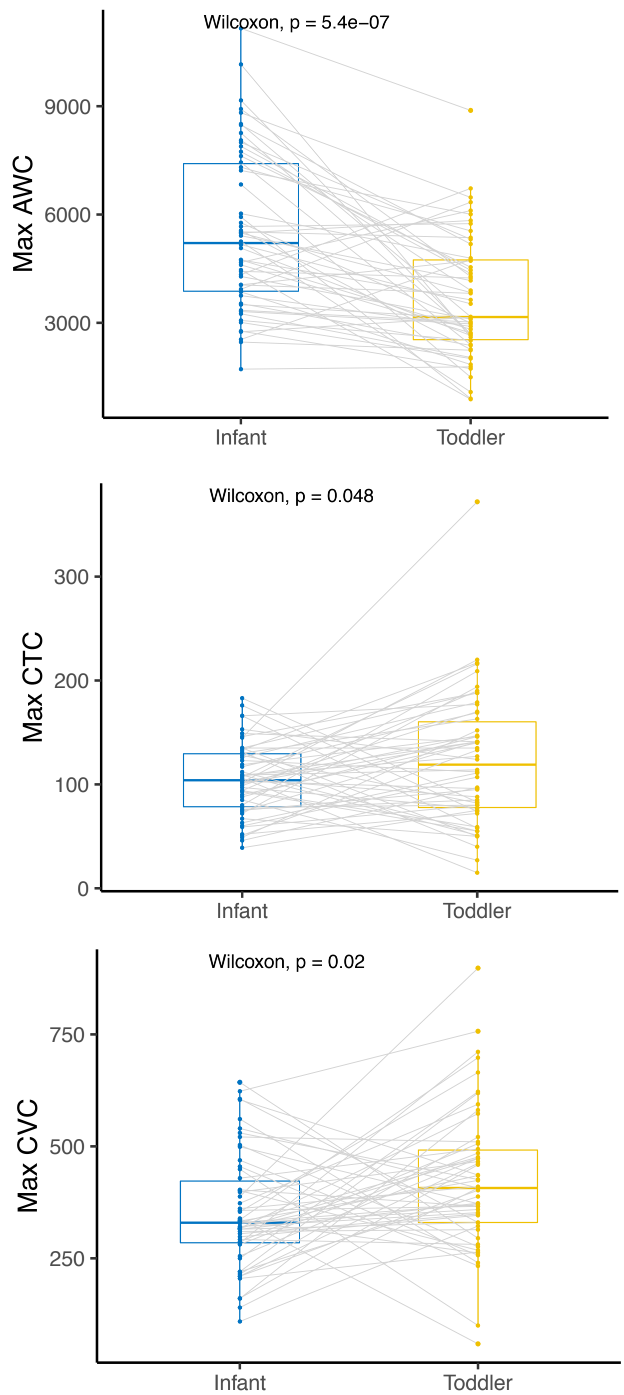


**Figure 3.1:** Adult word count, max turn count, and child vocalization count for the infant and toddler LENA recordings. The Infant and LENA Toddler recordings are presented in blue and yellow, respectively. Individual observations across observations are paired using grey lines. Results of paired Wilcoxon signed-rank tests for each LENA measure are indicated at the top of each plot.

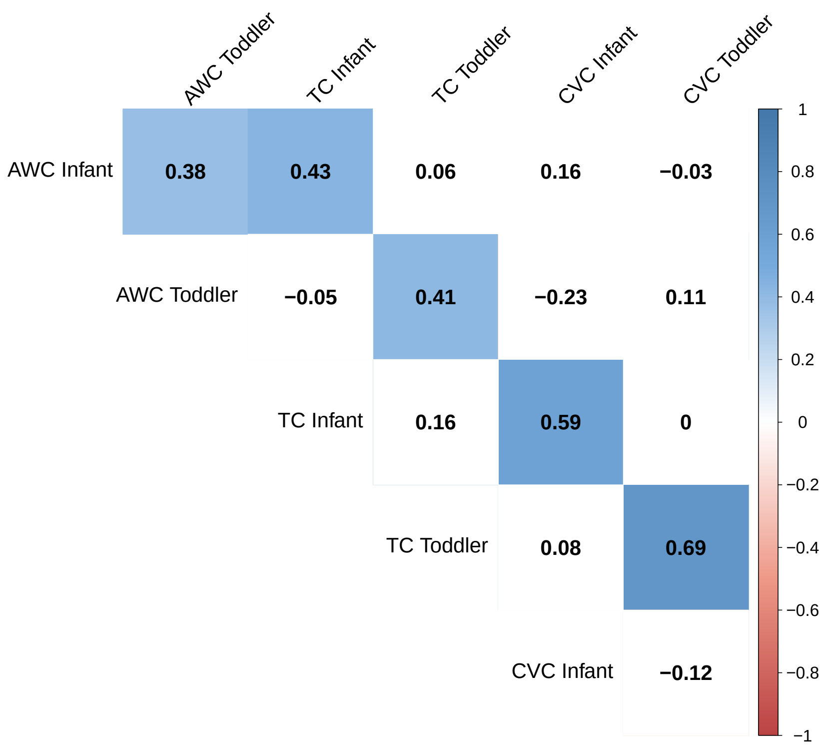


**Figure 3.2:** Correlation matrix for LENA Infant and LENA Toddler measurements including maximum adult word count (AWC), turn count (TC) and child vocalisations (CVC) per hour. Coloured squares display significant correlations. Positive correlations are in blue and negative correlations in red. Correlation coefficients are indicated inside each cell.

SES in subsequent models relating ELP and LENA AWC nor LENA CTC measures.

### 3.3.2 Language Input in Relation Language Processing

To examine longitudinal relations between the LENA and language processing measures, we ran several regression analyses with LENA input measures, AWC and CTC, predicting ELP 18 language processing measures or predicting ELP 30 language processing measures. All our models were fit with the LENA input measure (e.g. AWC or CTC) for both time points added as an aggregate fixed predictor of the relevant ELP measure. This simplified our models and allowed us to measure effects of each LENA observation individually without taking into account interactions between both LENA observations. We call this the "aggregated LENA". A linear model run with the lm function of the R package (R Core Team, 2017) was used for ELP measures that contain only one or two levels such as Reinforcement, Novelty bias, Referent Selection and Retention. Variables with two levels were scaled and centered. There was no random structure since we only had one or two observations per participant.

For ELP measures containing several levels, such as Word Difficulty and Word Type for ELP Comprehension, we used mixed effects models, particularly glmmTMB (Brooks Mollie et al., 2017) run with the R package (R Core Team, 2017). In this model, we set Word Type and Word Difficulty as main fixed effects, without interactions. This allowed us to control for Word Type and Word Difficulty in our model, while still looking at the effect of ELP Comprehension overall. We did not look at the possible interactions with Word Type and Word Difficulty for two reasons: 1) the sample size is small for analysis of this complexity, and 2) when assessing the best model fit using Anova and comparing the Akaike's Information Criterion (AIC; Wagenmakers & Farrell, 2004), a simpler model fit our data significantly better with a smaller AIC.

In the random effects structure, a random intercept was nested within participant, as well as an interaction between participant and word difficulty. This allowed each participant a random intercept and accounted for individual effects of participant interacting with word difficulty, for a maximally-specified model. The best random structure for the model was also assessed using AIC and the ANOVA tests comparing a model with a random structure including word type versus a structure including word difficulty. In this case, both models had the same AIC and the ANOVA test was not significant. However, a model including word type in the random structure detected quantile deviations when assessing the model residuals versus the model predictions. This did not occur when using word difficulty in the random structure, indicating that it was a better model. We set the model family to Gaussian because the proportion data was normally distributed, and thus it is expected to have a linear effect. Finally, the effect of each parameter in the models was assessed with an F test (ANOVA function from the R package), which tests whether the model terms are significant. Thus, we report F test results for our models. For each of the ELP measures (our predicted variable) we report models examining the influence of AWC, and separately CTC, on 18 and then 30 month ELP performance.

**LENA Input and ELP Reinforcement**

To measure the relationships between AWC and ELP Reinforcement at 18 months, we used a linear model with the aggregated LENA AWC (AWC Infant and AWC Toddler) predicting ELP Reinforcement. The effect of each parameter in the model was assessed with an F test that showed no significant relationships between the variables (see Table 3.2). The same model predicting ELP Reinforcement at 30 months, did not show any significant relationships either (see F test values on Table 3.3).

To assess the relationship between CTC and ELP Reinforcement at 18-months, looking proportions to the target for ELP Reinforcement were fit with the aggregated LENA CTC (CTC Infant and CTC Toddler) as predictors and main fixed effects. The effect of each parameter in the model was assessed with an F test that revealed a main effect of CTC Infant and a main effect of CTC Toddler (see Table 3.12). As can be seen in Figure 3.3, the number of conversational turns in infancy was negatively associated with looks to the target on ELP Reinforcement measures. By contrast, the number of conversational turns in toddlerhood was positively associated with looks to the target on ELP Reinforcement. This indicates that, when looking at highly familiar word recognition abilities, children benefit from conversational experiences at older ages.

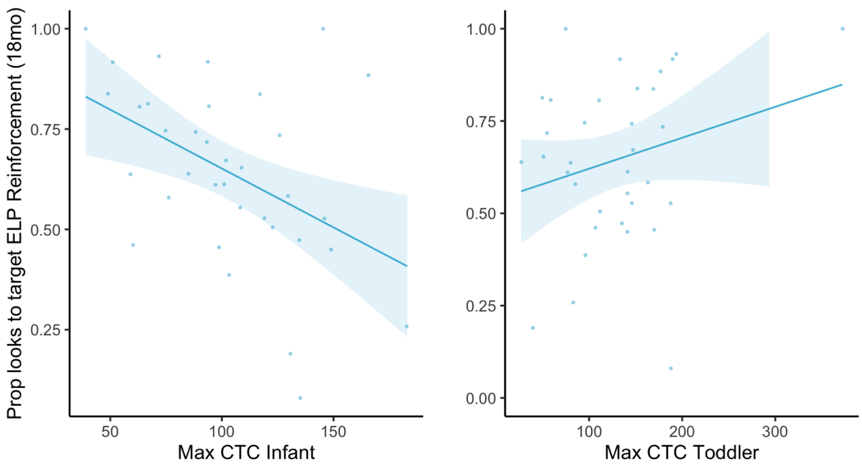


**Figure 3.3:** Relationships between CTC and ELP Reinforcement at 18 months. The left panel shows CTC in infancy, the right panel shows CTC in toddlerhood.

A linear model was fit with aggregated LENA CTC (CTC Infant and CTC Toddler) as fixed effects, predicting mean looking proportions to target on ELP Reinforcement at 30 months. The model showed no significant relationships (see F test Table 3.13).

### LENA Input and ELP Comprehension

To measure the relationship between AWC and ELP Comprehension at 18-months, looking proportions to target for ELP Comprehension split by word type (nouns, verbs and adjectives) and difficulty, were modelled using a mixed effects model. The model was fit with the aggregated LENA AWC (AWC Infant and AWC Toddler), Word Type and Word Difficulty as predictors. All the predictors were added as main effects (i.e., they did not interact with each other). In the random effects structure a random intercept was nested within participant, as well as an interaction between participant and word difficulty. The model revealed a main effect of Word Difficulty (see F test values on Table 3.4). This main effect indicates looking differences across word difficulties, which we have already reported in Chapter 2. There were no significant effects involving the input measures.

To measure the relationships between AWC and ELP Comprehension at 30-months, we used a mixed effects model with the same structure as the one used for ELP at 18 months. Looking proportions to the target for ELP Comprehension at 30 months were fit with the aggregated LENA AWC (AWC Infant and AWC Toddler), Word Type (noun, verbs and adjective) and Word Difficulty (easy, moderate and difficult) as predictors. The model revealed a main effect of Word Type (see F test values on Table 3.5). This main effect indicates looking differences across nouns, verbs and adjectives at 30-months, replicating our findings from Chapter 2. Again, there were no significant effects involving the input measures.

To assess the relationships between CTC and ELP Comprehension at 18-months, looking proportions to the target for ELP Comprehension split by word type (nouns, verbs and adjectives) and difficulty, were modelled using a mixed effects model fit with the aggregated LENA CTC (CTC Infant and CTC Toddler), Word Type, Word Difficulty as predictors. Fixed effects were set as fixed main effects. In the random effects structure a random intercept was nested within participant, as well as an interaction between participant and word difficulty. The effect of each parameter in the model was assessed with an F test that revealed a main effect of Word Difficulty (see Table 3.14). This main effect indicates overall looking differences across word difficulties, which we have also reported in Chapter 2. There were no significant effects involving the input measures.

To investigate the relationship between CTC and ELP Comprehension at 30 months, we used a mixed effects model. Mean looking proportions to target were fit with the aggregated LENA CTC (CTC Infant and CTC Toddler), main effects of Word Type and Word Difficulty as predictors. In the random effects structure a random intercept was nested within participant, as well as an interaction between participant and word difficulty. The model showed a main effect of Word Type and a main effect of CTC Infant (see F values on Table 3.15). As can be seen in Figure 3.4, children who experienced more conversational turns in infancy, looked more to the target on ELP Comprehension trials at 30 months of age. The main effect of word type has been reported in Chapter 2.

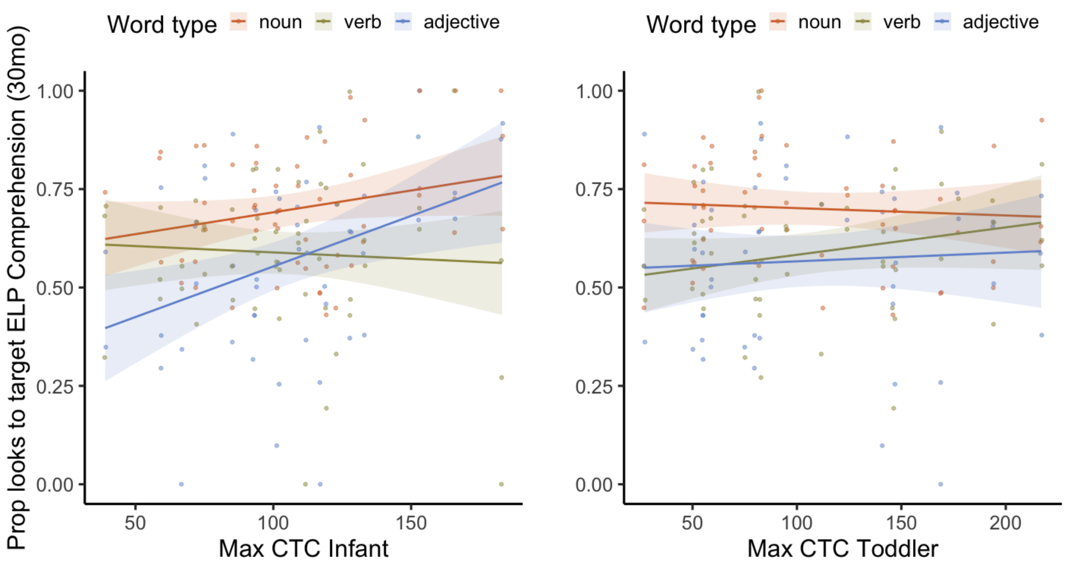


**Figure 3.4:** Relationships between CTC and Proportion looking to target on Comprehension ELP trials at 30 months. The left panel shows CTC in infancy, the right panel shows CTC in toddlerhood. Different colours indicate different word types.

**LENA Input and ELP Novelty Bias**

To measure the relationships between AWC and ELP Novelty Biases at 18-months, looking proportions to the novel item (i.e., ELP Novelty Bias) were fit with the aggregated LENA AWC (AWC Infant and AWC Toddler) and Familiar Image Type (familiar images that also appeared in ELP Comprehension, versus highly familiar images that also appeared in ELP Reinforcement) as predictors in a linear model. Results are reported in Table 3.6. The model showed a main effect of Familiar Image Type. The role of familiarity in novelty biases has already been reported in Chapter 2. There were no significant effects involving the input measures.

We used a linear model to measure relationships between AWC and Novelty Biases on language processing measures at 30-months. Looking proportions to the novel item were fit with aggregated LENA AWC (AWC Infant and AWC Toddler) and Familiar Image Type as fixed effects. The model did not show any significant relationships (see F test results in Table 3.7).

To assess the relationships between CTC and ELP Novelty Biases at 18-months, looking proportions to the novel image were fit with the aggregated LENA CTC (CTC Infancy and CTC Toddler) and Familiar Image Type as predictors. The effect of each parameter in the model was assessed with an F test (see Table 3.16). The model showed a main effect of Familiar Image Type. The role of familiarity in novelty biases has already been reported in Chapter 2. There were no significant effects involving the input measures.

A linear model was used to investigate possible relationships between CTC and ELP Novelty Bias measures at 30 months. Mean looking proportions to the novel image were fit with the aggregated LENA CTC (CTC Infant and CTC Toddler) and Familiar Image Type as predictors. The effect of each parameter in the model was assessed with an F test (see Table 3.17), which revealed no significant relationships.

### LENA Input and ELP Referent Selection

We used a linear model to assess the relationships between AWC and ELP Referent Selection at 18-months. Looking proportions to the target for ELP Referent Selection at 18-months were fit with the aggregated LENA AWC (AWC Infant and AWC Toddler) and Word Type (novel versus familiar noun) as predictors. The effect of each parameter in the model was assessed with an F test, showing a positive main effect of AWC Infant (see Table 3.8). Results are plotted in Figure 3.5: more adult input in infancy led to more overall looks to the target in ELP Referent Selection trials. This was the case when the target was a familiar noun and when the target was a novel word referring to a novel image.

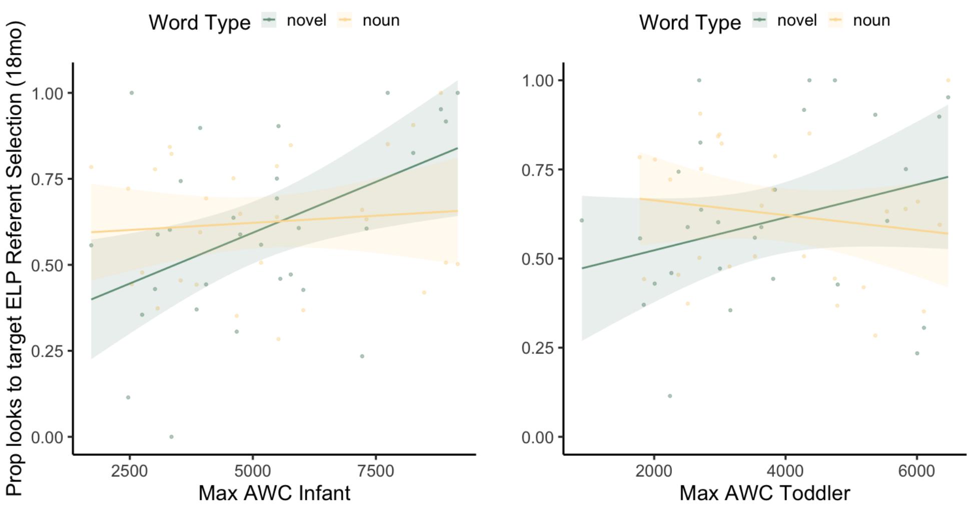


**Figure 3.5:** Relationships between AWC and ELP Referent Selection at 18 months. The left panel shows AWC in infancy, the right panel shows AWC in toddlerhood. Different colours indicate looks to target when that was novel, versus when it was a familiar noun.

A linear model was used to assess the relationships between AWC and ELP Referent Selection at 30 months. Looking proportions to the target on ELP

Referent Selection trials were fit with the aggregated LENA AWC (AWC Infant and AWC Toddler). In this case, due to the small sample size, we did not have data to account for looks to the target when this was novel; thus, this model only includes looks to the target when this was a familiar noun. Thus, Word Type (novel versus familiar) was removed from the fixed effects structure. This differs from the model used for Referent Selection at ELP 18-months. The effect of each parameter in the model was assessed with an F test. No relationships were found (see Table 3.9).

To measure relationships between CTC and ELP Referent Selection at 18-months, we ran a linear model. Looking proportions to the target for ELP Referent Selection trials were fit with the aggregated LENA CTC (CTC Infant and CTC Toddler) and Word Type (novel versus familiar noun). The effect of each parameter in the model was assessed with an F test (see Table 3.18), which revealed a main effect of CTC Infant. As can be seen in Figure 3.6, more conversational turns, particularly in infancy, led to more looks to the target for both novel words and familiar nouns at 18 months.

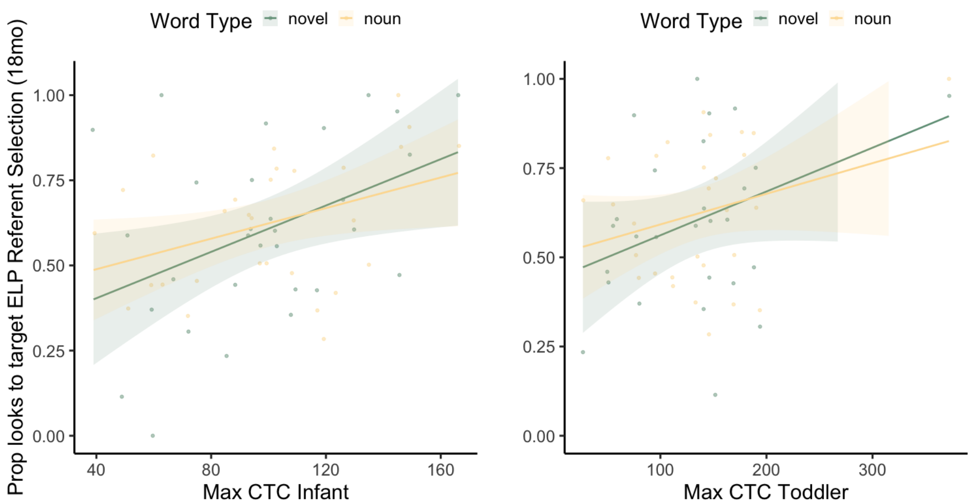


**Figure 3.6:** Relationships between CTC and ELP Referent Selection at 18 months. The left panel shows CTC in infancy, the right panel shows CTC in toddlerhood. Different colours indicate looks to target when that was novel, versus when it was a familiar noun.

A linear model was used to assess relationships between CTC and ELP Referent Selection at 30 months. The aggregated LENA CTC (CTC Infant and CTC Toddler) was set as fixed effects, predicting mean looking proportions to target on ELP Referent Selection trials. In this model, Target Type (novel versus familiar) could not be added to the fixed effects structure due to missing data for looks to target when this was novel. Thus, we only predict looks to the target when this was a familiar noun (in the context of a novel image). The effect of each parameter in the model was assessed with an F test. There were no significant relationships (see Table 3.19).

**LENA Input and ELP Retention**

Two linear models with the same structure were used to measure the relationships between AWC and ELP Retention at 18 months and at 30 months. Mean proportion of looks to the target on ELP Retention trials (either at 18 or 30 months) were fit with the aggregated LENA AWC (AWC Infant and AWC Toddler) as predictors. Neither model showed significant relationships (see Table 3.10 for the F values of the model predicting ELP Retention at 18 months, and Table 3.11 for the model using ELP Retention at 30 months).

To measure possible relationships between CTC and ELP Retention at 18 months, we used a linear model. Looking proportions to the target on ELP Retention trials were fit with the aggregated LENA CTC (CTC Infant and CTC Toddler) as predictors. We assessed the effect of each parameter in the model with an F test, which revealed no significant relationships (see Table 3.20).

Relationships between CTC and ELP Retention at 30 months were assessed using the same linear model as in ELP 18-months. Mean looking proportions to target were fit with the aggregated LENA CTC (CTC Infant and CTC Toddler) as predictors. An F test on the model's parameters showed no significant relationships (see Table 3.21).

## 3.4 Discussion

The goal of this study was to explore whether children's language input influenced their language processing abilities. To that aim, we collected approximately 16 hours of day-long home recordings across three days at two time points – in infancy and during toddlerhood. We used the LENA system to compute estimates of the amount of adult words and conversations turns that children were exposed to at home. When children were both 18 and 30 months, they did the ELP task that measured children's speed of word processing, word comprehension, novelty biases, referent selection abilities and retention of new words based on children's looking behaviour. We used linear models to measure the strength of the relationships between language input and ELP language outcome measures. Overall our results show stronger effects for language input measured in infancy and conversational turns.

### 3.4.1 Language Input and Online Comprehension abilities

Our findings showed negative relationships between the number of conversational turns in infancy and performance on ELP Reinforcement measured at 18 months. By contrast, the number of conversational turns in toddlerhood was positively associated with looks to the target on ELP Reinforcement at 18 months. Thus, children's abilities to recognise the highly familiar words used on the ELP Reinforcement trials, benefited from having heard more turns as toddlers but not as infants. Previous literature has reported that children might benefit from different aspects of their input at different ages (e.g., Rowe, 2012). Quantity of input, such as the amount of adult words, might be more relevant in infancy, because it might help children in the first steps of the language learning process. In contrast, conversational turns, which contain qualitative aspects of the language input, might be more helpful later in development, when children have more experience with language and are able to benefit from the richness of that speech. This hypothesis is a possible explanation for our pattern of results that shows that children's abilities to recognise highly familiar words were enhanced with larger amounts of input quality during toddlerhood, but not during infancy, when conversational turns might be less meaningful. Nevertheless, we found a positive relationship between conversational turns in infancy and ELP Comprehension measures at 30 months. Comprehension trials included nouns, verbs and adjectives of different difficulties. Thus, our data suggests that greater conversational experience in infancy have a long-term impact in children's word recognition abilities.

An alternative explanation for this complex pattern of results might lie on the nature of conversational experience in infancy versus in toddlerhood. Our infant group included children who were 6.91 months on average. It is possible that conversational turns at that age relay less on the linguistic aspect of conversational experience, and more on the rich social context in which the interaction occurs. A recent longitudinal study using LENA, tested the developmental relationship between conversational turn-taking and vocabulary growth in English-acquiring children between 9 and 24 months (Donnelly & Kidd, 2020). The study showed a bidirectional relationship between conversational turns and vocabulary growth across early development, controlling for the amount of words in children's environments. Thus, it is possible that early social interaction is beneficial for long term language abilities via vocabulary growth. Greater social interactions though conversational experience in infancy, might have boosted children's conversational experience and vocabulary growth over the first years of development, which resulted in better word recognition abilities later on. Our data might be able to capture the long term effects of early social interaction on language abilities at 30 months. The long term benefit of early social interactions might have been hard to capture in ELP Reinforcement and Comprehension measures at 18 months and thus, we find negative or non significant relationships between conversational experience in infancy and performance on ELP word comprehension measures at 18 months.

We do find however, a positive relationship between conversational turns in toddlerhood and ELP Reinforcement at 18 months. This might be evidence that what LENA is measuring as conversational turns in infancy (at approximately 6 months) is not the same that is measuring in toddlerhood (at approximately 20 months). Early in infancy, conversational turns estimated with LENA might be more related to the social interaction between caregiver and child, whereas estimates of conversational turns in toddlerhood might contain more linguistic features and they might be more related to infants interactions with objects. This could make turn estimates early in infancy a noisy measure. In fact, LENA turn count measures have been reported to be their less accurate estimate (Ferjan Ramírez, Hippe, & Kuhl, 2021; Cristia et al., 2021). Thus, it is possible that LENA is miss-counting turns, particularly early in development.

If this is the case, then it makes sense that conversational turns in toddlers, containing a rich input, is associated with better word recognition abilities of highly familiar words measured at a similar point in time. Note that both of those measures, LENA toddler and ELP at 18, were collected at a very close point in time. These hypothesis cannot be confirmed with our current data.

A possibility might be to look at what characterises conversational tuns measured by LENA in infancy versus during toddlerhood. Another option could be to use meditation analysis to assess if vocabulary growth mediates the relationship between early conversational experience and later online comprehension abilities.

### 3.4.2 Language Input and Referent Selection abilities

Our results showed positives relationships between both the amount of adult words and conversational turns during infancy and referent selection abilities at 18-month-old. This is in line with previous literature showing that both quantity and quality of language input early in development is associated with language abilities later on (Rowe, 2012). It makes sense that quantity of adult words is important, particularly in situations where learning can be difficult. In ELP Referent Selection trials, children had to use disambiguation to correctly map the novel word into the novel object. This requires a strong knowledge of the familiar image to be able to exclude it as possible referent and map the novel word into the novel image, in other words, larger vocabularies. Larger quantities of adult words might have help children gain a reasonable knowledge of some words because they have appeared many times in their input. Likewise, conversational experience, either in the form of social interaction or with a high linguistic exchange, might have presented the child with situations in which a novel word appeared in the context of a novel object, for example during playing sessions. In our data however, early input also help children to select the familiar object, when this was the target, in the context of a novel one. This skills might have been enhanced by both quantity and quality of language input during infancy. In language acquisition literature, there are several studies showing evidence of relationships between the input, child vocabulary size, the speed of word processing and word learning. Children who have been exposed to richer input possess a larger vocabulary and process words quicker than their peers. In turn, this affects the speed with which children learn new words and, ultimately, has an effect on the size of their vocabulary in later years (Weisleder & Fernald, 2013; Hurtado et al., 2008b). In fact, the advantages of faster processing speed in young children can still be seen several years after (Marchman & Fernald, 2008). A study that measured quantity and quality of linguistic input independent using computational models, showed that whereas quantity of input mattered initially to boost learning of new words, quality of input, indexed by lexical diversity, was ultimately more beneficial for the model to learn new words (Jones & Rowland, 2017). In our data both quantity and quality were not independent constructs because conversational turns were computed based on adults words and child vocalisations that occurred in speech within less than a 5 second pause. This makes it difficult to isolate the influence of each variable on our data and this could be the reason why we see that both adult input and conversational in infancy matter for children's disambiguation abilities. Ultimately, our relationships might indicate that both quantity and quality early in development benefit children's abilities to find the correct referent in a word learning task.

Most of the relationships we found between language input and language processes were in Infancy and they did not hold for our 30-month ELP measurements. A possibility is that the small sample size at that time point makes conclusions difficult. Another explanation is that LENA measures are too distant in time to show an effect on the ELP 30 months lanaguge processing measures. This study is part of a larger longitudinal project that has collected LENA measures closer in time to ELP 30 months. Further analysis could help understand if input measured at a closer time point have an effect on ELP performance at 30 months.

From the relations we did find, it appears that children's own participation in dyadic interactions, as measured by child turn count is more predictive of later processing than adult word input. This might reflect more qualitative proprieties of the input they are exposed to as well as children's own aptitudes which might also be beneficial long term. However, since turn estimates from the LENA have been reported to be less accurate than other input estimates. Future analysis could look into the content of conversational turns early in infancy using coding schemes (e.g., ACLEW Soderstrom et al., 2020) or use new automatised alternatives to the LENA (e.g., Räsänen, Seshadri, Lavechin, Cristia, & Casillas, 2021).

This study is limited to a western sample of children learning British English. We do not know how our findings generalise to other populations and languages. It is possible that other contextual variables play a role, as well as the family structure. Children living in larger households might be exposed to more overheard speech. We know that children benefit from speed directed to the child. So it might be challenging for children to learn language mostly based on speech that was not directed to them.

## 3.5 Conclusion

Multiple studies have highlighted the role of language input as a key variable that has an impact on children's language development. Here we explored whether children's language input influenced their language processing abilities at two time points.

To that aim, we collected day-long home recordings in infancy and during toddlerhood. We used the LENA system to compute estimates of the amount of adult words and conversations turns that children were exposed to at home. The same children did the ELP task when they were 18 and 30 months old. ELP measured children's speed of word processing, word comprehension, novelty biases, referent selection abilities and retention of new words based on children's looking behaviour. Most of the relationships we found between language input and language processes were in Infancy with stronger effects conversational turns, particularly for measures related to comprehension and word learning. This highlights the role of early conversational experience in relation to lexical skills as well as the ability to learn new words.

Based on this findings, future studies could look at the relationships between conversational experience and language outcomes across development to measure the long term consequences of early experiences for children's language development.

## 3.6 Significance Tables

**Table 3.2:** Regression results for AWC and ELP Reinforcement at 18 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 14.64 | 1.00 | 323.86 | 0.00 | *** |
| AWCinfant | 0.14 | 1.00 | 3.18 | 0.08 | . |
| AWCtoddler | 0.17 | 1.00 | 3.65 | 0.07 | . |
| Residuals | 1.45 | 32.00 | | | |

*Note. Main fixed effects are displayed including AWC (adult word count) Infant ((M = 6.91 months) and AWC Toddler (M = 20.25 months). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.3:** Regression results for AWC and ELP Reinforcement at 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| | | | | | |
|---|---|---|---|---|---|
| (Intercept) | 11.82 | 1.00 | 493.25 | 0.00 | *** |
| AWCinfant | 0.00 | 1.00 | 0.10 | 0.75 | |
| AWCtoddler | 0.01 | 1.00 | 0.44 | 0.52 | |
| Residuals | 0.41 | 17.00 | | | |

*Note. Main fixed effects are displayed including AWC (adult word count) Infant ((M = 6.91 months) and AWC Toddler (M = 20.25 months). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.4:** Regression results for AWC and ELP Comprehension at 18 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 141.41 | 1.00 | 0.00 | *** |
| AWCinfant | 1.16 | 1.00 | 0.28 | |
| AWCtoddler | 0.88 | 1.00 | 0.35 | |
| WordDiff | 6.19 | 2.00 | 0.04 | * |
| WordType | 5.74 | 2.00 | 0.06 | . |

*Note. Main fixed effects are displayed including AWC (adult word count) Infant ((M = 6.91 months), AWC Toddler (M = 20.25 months), Word Difficulty (easy, moderate and difficult) and Word Type (nouns, verbs and adjectives). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.5:** Regression results for AWC and ELP Comprehension at 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 204.57 | 1.00 | 0.00 | *** |
| AWCinfant | 0.05 | 1.00 | 0.82 | |
| AWCtoddler | 0.93 | 1.00 | 0.33 | |
| WordDiff | 2.75 | 2.00 | 0.25 | |
| WordType | 16.83 | 2.00 | 0.00 | *** |

*Note. Main fixed effects are displayed including AWC (adult word count) Infant ((M = 6.91 months), AWC Toddler (M = 20.25 months), Word Difficulty (easy, moderate and difficult) and Word Type (nouns, verbs and adjectives). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.6:** Regression results for AWC and ELP Novelty Bias at 18 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 16.91 | 1.00 | 362.71 | 0.00 | *** |
| AWCinfant | 0.00 | 1.00 | 0.00 | 0.99 | |
| AWCtoddler | 0.03 | 1.00 | 0.69 | 0.41 | |
| FamiliarImageType | 0.37 | 1.00 | 7.98 | 0.01 | ** |
| AWCinfant:FamiliarImageType | 0.13 | 1.00 | 2.80 | 0.10 | |
| AWCtoddler:FamiliarImageType | 0.08 | 1.00 | 1.81 | 0.19 | |
| Residuals | 2.00 | 43.00 | | | |

*Note. Main fixed effects are displayed including AWC (adult word count) Infant ((M = 6.91 months), AWC Toddler (M = 20.25 months) and Familiar Image Type (familiar images that appeared in ELP Comprehension, versus highly familiar images that appeared in ELP Reinforcement). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.7:** Regression results for AWC and ELP Novelty Bias at 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 14.20 | 1.00 | 320.83 | 0.00 | *** |
| AWCinfant | 0.00 | 1.00 | 0.08 | 0.78 | |
| AWCtoddler | 0.03 | 1.00 | 0.67 | 0.42 | |
| FamiliarImageType | 0.01 | 1.00 | 0.21 | 0.65 | |
| AWCinfant:FamiliarImageType | 0.00 | 1.00 | 0.11 | 0.75 | |
| AWCtoddler:FamiliarImageType | 0.00 | 1.00 | 0.01 | 0.94 | |
| Residuals | 1.37 | 31.00 | | | |

*Note. Main fixed effects are displayed including AWC (adult word count) Infant ((M = 6.91 months), AWC Toddler (M = 20.25 months) and Familiar Image Type (familiar images that appeared in ELP Comprehension, versus highly familiar images that appeared in ELP Reinforcement). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.8:** Regression results for AWC and ELP Referent Selection at 18 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 22.32 | 1.00 | 463.62 | 0.00 | *** |
| AWCinfant | 0.28 | 1.00 | 5.88 | 0.02 | * |
| AWCtoddler | 0.00 | 1.00 | 0.09 | 0.77 | |
| WordType | 0.01 | 1.00 | 0.16 | 0.69 | |
| AWCinfant:WordType | 0.07 | 1.00 | 1.45 | 0.23 | |
| AWCtoddler:WordType | 0.08 | 1.00 | 1.60 | 0.21 | |
| Residuals | 2.55 | 53.00 | | | |

*Note. Fixed effects are displayed including AWC (adult word count) Infant ((M = 6.91 months), AWC Toddler (M = 20.25 months) and Word Type (familiar noun versus novel). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.9:** Regression results for AWC and ELP Referent Selection at 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 9.74 | 1.00 | 496.52 | 0.00 | *** |
| AWCinfant | 0.00 | 1.00 | 0.10 | 0.75 | |
| AWCtoddler | 0.00 | 1.00 | 0.00 | 0.95 | |
| Residuals | 0.31 | 16.00 | | | |

*Note. Fixed effects are displayed including AWC (adult word count) Infant ((M = 6.91 months) and CTC Toddler (M = 20.25 months), only looks to target when this was a familiar noun are analysed. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.10:** Regression results for AWC and ELP Retention at 18 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 8.07 | 1.00 | 148.60 | 0.00 | *** |
| AWCinfant | 0.02 | 1.00 | 0.29 | 0.59 | |
| AWCtoddler | 0.04 | 1.00 | 0.77 | 0.39 | |
| Residuals | 1.30 | 24.00 | | | |

*Note. Fixed effects are displayed including AWC (adult word count) Infant ((M = 6.91 months) and AWC Toddler (M = 20.25 months). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.11:** Regression results for AWC and ELP Retention at 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 4.61 | 1.00 | 59.58 | 0.00 | *** |
| AWCinfant | 0.00 | 1.00 | 0.00 | 0.96 | |
| AWCtoddler | 0.09 | 1.00 | 1.18 | 0.30 | |
| Residuals | 1.01 | 13.00 | | | |

*Note. Fixed effects are displayed including AWC (adult word count) Infant ((M = 6.91 months) and AWC Toddler (M = 20.25 months). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.12:** Regression results for CTC and ELP Reinforcement at 18 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 14.78 | 1.00 | 462.51 | 0.00 | *** |
| CTCinfant | 0.55 | 1.00 | 17.27 | 0.00 | *** |
| CTCtoddler | 0.30 | 1.00 | 9.38 | 0.00 | ** |
| Residuals | 1.02 | 32.00 | | | |

*Note. Main fixed effects are displayed including CTC (conversational turn count) Infant ((M = 6.91 months) and CTC Toddler (M = 20.25 months). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.13:** Regression results for CTC and ELP Reinforcement at 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 11.85 | 1.00 | 489.27 | 0.00 | *** |
| CTCinfant | 0.00 | 1.00 | 0.14 | 0.71 | |
| CTCtoddler | 0.00 | 1.00 | 0.07 | 0.79 | |
| Residuals | 0.41 | 17.00 | | | |

*Note. Main fixed effects are displayed including CTC (adult word count) Infant ((M = 6.91 months) and CTC Toddler (M = 20.25 months). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.14:** Regression results for CTC and ELP Comprehension at 18 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 142.85 | 1.00 | 0.00 | *** |
| CTCinfant | 2.49 | 1.00 | 0.12 | |
| CTCtoddler | 0.09 | 1.00 | 0.76 | |
| WordDiff | 6.20 | 2.00 | 0.04 | * |
| WordType | 5.73 | 2.00 | 0.06 | . |

*Note. Main fixed effects are displayed including CTC (conversational turn count) Infant ((M = 6.91 months), CTC Toddler (M = 20.25 months), Word Difficulty (easy, moderate and difficult) and Word Type (nouns, verbs and adjectives). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.15:** Regression results for CTC and ELP Comprehension at 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 216.75 | 1.00 | 0.00 | *** |
| CTCinfant | 4.20 | 1.00 | 0.04 | * |
| CTCtoddler | 0.07 | 1.00 | 0.80 | |
| WordDiff | 2.89 | 2.00 | 0.23 | |
| WordType | 16.20 | 2.00 | 0.00 | *** |

*Note. Main fixed effects are displayed including CTC (conversational turn count) Infant ((M = 6.91 months), CTC Toddler (M = 20.25 months), Word Difficulty (easy, moderate and difficult) and Word Type (nouns, verbs and adjectives). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.16:** Regression results for CTC and ELP Novelty Bias at 18 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 15.11 | 1.00 | 334.97 | 0.00 | *** |
| CTCinfant | 0.15 | 1.00 | 3.28 | 0.08 | . |
| CTCtoddler | 0.13 | 1.00 | 2.87 | 0.10 | . |
| FamiliarImageType | 0.43 | 1.00 | 9.64 | 0.00 | ** |
| CTCinfant:FamiliarImageType | 0.02 | 1.00 | 0.38 | 0.54 | |
| CTCtoddler:FamiliarImageType | 0.03 | 1.00 | 0.73 | 0.40 | |
| Residuals | 1.94 | 43.00 | | | |

*Note. Main fixed effects are displayed including CTC (conversational turn count) Infant ((M = 6.91 months), CTC Toddler (M = 20.25 months) and Familiar Image type (familiar images that appeared in ELP Comprehension, versus highly familiar images that appeared in ELP Reinforcement). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.17:** Regression results for CTC and ELP Novelty Bias at 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 14.15 | 1.00 | 320.37 | 0.00 | *** |
| CTCinfant | 0.01 | 1.00 | 0.21 | 0.65 | |
| CTCtoddler | 0.01 | 1.00 | 0.24 | 0.63 | |
| FamiliarImageType | 0.02 | 1.00 | 0.40 | 0.53 | |
| CTCinfant:FamiliarImageType | 0.02 | 1.00 | 0.39 | 0.54 | |
| CTCtoddler:FamiliarImageType | 0.01 | 1.00 | 0.12 | 0.73 | |
| Residuals | 1.37 | 31.00 | | | |

*Note. Main fixed effects are displayed including CTC (conversational turn count) Infant ((M = 6.91 months), CTC Toddler (M = 20.25 months) and Familiar Image type (familiar images that appeared in ELP Comprehension, versus highly familiar images that appeared in ELP Reinforcement). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.18:** Regression results for CTC and ELP Referent Selection at 18 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 21.90 | 1.00 | 459.66 | 0.00 | *** |
| CTCinfant | 0.27 | 1.00 | 5.70 | 0.02 | * |
| CTCtoddler | 0.04 | 1.00 | 0.94 | 0.34 | |
| WordType | 0.00 | 1.00 | 0.05 | 0.82 | |
| CTCinfant:WordType | 0.01 | 1.00 | 0.25 | 0.62 | |
| CTCtoddler:WordType | 0.00 | 1.00 | 0.01 | 0.90 | |
| Residuals | 2.53 | 53.00 | | | |

*Note. Fixed effects are displayed including CTC (conversational turn count) Infant ((M = 6.91 months), CTC Toddler (M = 20.25 months) and Word Type (familiar noun versus novel). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.19:** Regression results for CTC and ELP Referent Selection at 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 9.72 | 1.00 | 530.17 | 0.00 | *** |
| CTCinfant | 0.00 | 1.00 | 0.08 | 0.79 | |
| CTCtoddler | 0.02 | 1.00 | 1.23 | 0.28 | |
| Residuals | 0.29 | 16.00 | | | |

*Note. Fixed effects are displayed including CTC (conversational turn count) Infant ((M = 6.91 months) and CTC Toddler (M = 20.25 months), only looks to target when this was a familiar noun are analysed. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.20:** Regression results for CTC and ELP Retention at 18 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 8.06 | 1.00 | 151.81 | 0.00 | *** |
| CTCinfant | 0.03 | 1.00 | 0.56 | 0.46 | |
| CTCtoddler | 0.15 | 1.00 | 2.90 | 0.10 | |
| Residuals | 1.27 | 24.00 | | | |

*Note. Fixed effects are displayed including CTC (conversational turn count) Infant ((M = 6.91 months) and CTC Toddler (M = 20.25 months). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 3.21:** Regression results for CTC and ELP Retention at 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 4.54 | 1.00 | 64.97 | 0.00 | *** |
| CTCinfant | 0.03 | 1.00 | 0.41 | 0.53 | |
| CTCtoddler | 0.18 | 1.00 | 2.54 | 0.14 | |
| Residuals | 0.91 | 13.00 | | | |

*Note. Fixed effects are displayed including CTC (conversational turn count) Infant ((M = 6.91 months) and CTC Toddler (M = 20.25 months). Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

# Chapter 4

# Adapting the Early Language Processing Task to a rural Indian Context

## 4.1 Introduction

It is well established in literature that early language skills are predictive of later language and cognitive abilities. Early language has been related to overall intellectual ability (e.g., Feldman et al., 2005), the development of executive functions (e.g., Wade et al., 2014), literacy outcomes (J. Lee, 2011; Duff et al., 2015) and academic success (e.g., Agostin & Bain, 1997). Specific early language abilities, such as speech perception, segmentation, and word recognition skills measured in young children predict vocabulary measures in the second and third years of life (see Cristia et al., 2014 for a systematic review). This suggests that early language abilities support cognitive skill and intellectual functioning from early in development. Studying the mechanisms that support early word learning is crucial to better understand how these early processes relate to later language development and cognitive outcomes.

Language development is also influenced by children's social context (Hoff, 2006). One of the variables that has been associated with cognitive development, including language skill, is early adversity. Studies show that children experiencing early adversity, such as nutritional deficits or limited or poor-quality adult input, are at a high risk of delays in their cognitive development (Leroy, Gadsden, & Guijarro, 2012; L. C. Fernald et al., 2012). This represents a high percentage of children in low and middle income countries. It is estimated that 250 million children (about a 43%) in low and middle income countries fail to reach their developmental potential due to early adversity (Black et al., 2017). Poverty and early adversities significantly impact development, accentuating the risk of poor socioeconomic outcomes and contributing to a cycle of poverty. Associations between socioeconomic status (SES) and early growth faltering (i.e., stunting), memory (Wijeakumar, Kumar, Delgado Reyes, Tiwari, & Spencer, 2019), executive function (L. C. Fernald et al., 2012), brain development (Hackman & Farah, 2009; Noble et al., 2012) and language (Hart & Risley, 1995) have been well established in infancy and childhood and across the lifespan (Bradley & Corwyn, 2002; Brooks-Gunn & Duncan, 1997; Kelly, Sacker, Del Bono, Francesconi, & Marmot, 2011).

When looking only at the relationships between SES and language abilities, the literature shows that differences in SES are strongly associated with variation in children's language outcomes. By the time they enter kindergarten, children from disadvantaged backgrounds differ substantially from their more advantaged peers in verbal and other cognitive abilities (Ramey & Ramey, 2004), disparities that are predictive of later academic success (V. E. Lee & Burkam, 2002). These SES differences in language proficiency can still be seen in adults indicating that SES differences in language skills are robust and cumulative, and expand across the lifetime (Pakulak & Neville, 2010). This is critical because SES differences in language abilities have been documented early in development. Already by 18 months of age, children show significant disparities in vocabulary and language processing efficiency based on SES differences. These differences across higher and lower SES families become critical for language development by 24 months, because there is evidence of a 6-month gap between SES groups in language processing skills (A. Fernald, Marchman, & Weisleder, 2013). Thus, it is clear that SES starts to influence language processing abilities early on, and if those disparities carry on, they can become critical for language skills in toddlerhood and still be present in adulthood. To boost children's language abilities, it is vital to study the relationships between contextual factors such as SES and language development, so we can implement early interventions that help children reach their full potential.

A challenge here is that most of our studies documenting relationships between children's context and language abilities come from western samples. Some of the studies looking at early adversity that we have previously reviewed document SES effects across low and middle income countries on children's cognitive abilities (e.g., L. C. Fernald et al., 2012), but they do not measure the effect of adversity on specific mechanisms that support early word learning. Thus, we do not know how findings generalise to other populations. This gap could be due to the fact that there are not many tasks that can be used cross-culturally, particularly tasks that measure language ability. This is because standardised measures of language ability are usually developed in western cultures; such measures can be difficult to translate to understudied languages and may not be culturally appropriate. For example, the Mullen Scales of Early Learning (Mullen & Others, 1995) at 2 years include questions about numbers and colours; such abstract concepts are unlikely to play a key role in lower-income, non-western contexts. Measuring early language becomes even more complex when attempting to asses language abilities in pre-verbal infants. In western cultures, word comprehension is usually measured using parental checklists, where parents are given a list of words and asked to report their child's understanding of those words. To use this approach across cultures, we need a validated parental vocabulary checklist for every language and cultural context. This is a daunting requirement. Moreover, parental checklists must be verbally administered to illiterate parents.

Some assessments have been developed that push through these challenges. One study developed a language assessment appropriate for use in the Wolof language and culture. This assessment included two measures based on caregiver report and it was used to measure language skill (language milestones achieved and vocabulary knowledge) in Wolof-learning infants and toddlers living in rural African villages. The authors assessed the psychometric properties and performance of two caregiver-report measures of Wolof language skill finding that both caregiver-report measures had good psychometric properties and displayed expected age and socioeconomic effects (Weber, Marchman, DIop, & Fernald, 2018). Even adaptations of assessments originally developed in the west have proved to be valid in very different populations. A recent study successfully adapted the Mullen Scales of Early Learning for use with infants in rural Gambia, including the language scale of this measure. Using children's scores on this adapted scale, the authors were able to examine cognitive development in the first 24-months of life as well as assess the association between cognitive performance and physical growth in children living in rural Gambia (Milosavljevic et al., 2019). These studies show that is it possible to adapt western measures to other languages and populations in a culturally relevant way. Still, caregiver reports or standardised scales do not provide direct measures of the child's language ability and thus, they do not assess the mechanisms involved in language processing abilities. As we can see, there is a need to develop direct tools that can be used across different populations and languages. This will help expand research to non-western samples, which will contribute to a better understanding of the mechanisms that support language development around the world. This will also help design interventions that boost children's language abilities early on.

The present project aims to contribute to the scarce literature measuring early language processing abilities in non-western samples, while using a direct, culturally-valid measure of children's language processing abilities. Our goal is to assess what language abilities children bring to the word learning problem. This study builds on Chapter 2 and extends it to an at-risk population in India. The aim of this project is to measure language processing abilities in children from India, particularly from Shivgarh, a rural village from Uttar Pradesh. The state of Uttar Pradesh is the most populous state in India and scores amongst the worst in terms of human development indicators. In rural Uttar Pradesh, in 2017 - 2018, the literacy rate among men was 80.5% and among women was 60.4%, this included persons aged 7 and above (data extracted from a report on education in India as part of *Key Indicators of Household Social Consumption on Education in India. NSS 75th Round*, 2018). The local dialect is called Awadhi, and it is reported to have 4 million native speakers in 2011 in India (*Language. India, States and Union Territories*, 2011). Awadhi belongs to the Indo-European languages, particularly the Indo-Aryan sub-family. It is generally viewed as a rural tongue, yet people in urban areas tend to speak a mixed form of Awadhi with standard Hindi, whereas in rural areas people speak only Awadhi. Education in rural areas in Uttar Pradesh is in Hindi but there is considerable epic literature written in Awadhi (for a detailed study of Awadhi language and its characteristics see Saksena, 1971).

To measure language processing abilities in children from Shivgarh learning Awadhi, we adapted the Early Processing (ELP) task to a different language, population and culture. The ELP task is an eye-tracking based task that is able to capture multiple language processes of individual children by combining different measures of language processing in a relatively short assessment lasting 15 - 20 minutes. This allows the researcher to examine how constellations of language processes influence one another. The ELP task is a two-image looking-based task, inspired by several well-established measures of language processes, some of which have been shown to be predictive of later language skill: speed of word processing based on work from A. Fernald et al. (1998), novelty bias, referent selection (or disambiguation), and retention of new words based on studies using the Reference Selection and Retention (RSR) task with either 2D or 3D images (Bion et al., 2013; Samuelson et al., 2017; Horst & Samuelson, 2008). ELP also incorporates an online measure of word comprehension which gives a direct measure of a child's vocabulary size or word comprehension abilities, which is based on the Computerised Comprehension Task (CCT) (Friend & Keplinger, 2008). We review this literature below as in Chapter 2 with an emphasis on studies using this measures in understudied languages, populations and across different SES. A second aim of this project is to capture the developmental trajectory of early word learning; thus, we administered the ELP to the same group of children at two different time points (at 18 and 30 months of age). With an eye towards large scale deployment in multiple populations, we wanted ELP to be a portable, efficient and adaptable looking based task.

The language processing measures integrated in the ELP task were specially selected because they are crucial elements of the task of learning a new word. For children to learn a word, they need to be able to extract a word from running speech, map it to the correct referent and remember that association later on. This process is influenced by the child's biases as well was experiences; vocabulary knowledge helps mapping the novel word onto the correct referent by excluding familiar referents in a process called disambiguation (Merriman & Schuster, 1991), and lexical processing skills may promote encoding the sentence context surrounding novel words, which can provide strong cues to their meanings (A. Fernald et al., 2006). Moreover, some of these processes have shown to be predictive of later language and cognitive skill.

### 4.1.1 Speed of Word Processing and Online Word Comprehension

A well known measure of language ability is speed of word processing (SoP), defined as how fast a child looks to a familiar image in response to a familiar spoken word when their first look was towards a distractor image, a behaviour usually assessed using the looking while learning (LWL) paradigm (A. Fernald et al., 1998, 2006). This measure is able to capture differences in children's word processing efficiency across different ages. In their seminal study, A. Fernald et al. (1998) examined the time course of word recognition in infants ages 15 to 24 months, finding that efficiency of verbal processing increases dramatically with age. Specifically, 15-month-old infants did not orient to the correct picture until after the target word was spoken, whereas 24-month-olds started shifting their gaze to the correct picture before the end of the spoken word. Speed of word processing is a particularly interesting measure because it has been associated with vocabulary skill and novel word learning abilities such that children with faster word recognition abilities have larger vocabulary sizes (A. Fernald et al., 2006), are better at learning new words (Lany, 2018), and show more accelerated growth in expressive vocabulary later on (A. Fernald et al., 2001; Marchman & Fernald, 2008). Furthermore, word processing abilities and vocabulary size during toddlerhood have been shown to be predictive of linguistic and cognitive skills at 8 years of age. Speed of word processing is also sensitive to contextual variables such as SES. Children from lower SES families show slower language processing efficiency at 18 months, and by 25 months, there is a 6-month gap between children from low and high SES (A. Fernald et al., 2013).

There is good evidence of relationships between early lexical processing efficiency and vocabulary size. However, because most of the studies reviewed relate speed of processing with vocabulary size using vocabulary size scores obtained though parental report, we do not know how the mechanisms involved with language processing and word comprehension relate to one another. Vocabulary checklists are powerful and well established tools that allow researchers to asses comprehension and production vocabulary sizes, however they do not tap directly into the cognitive mechanisms behind word comprehension abilities. Better understanding of infant's word comprehension is crucial because it provides the earliest window into children's understanding of word-referent relationships (Bates, 1993), and it is predictive of later language skills. In a study by Duff et al. (2015), pre-literacy vocabulary knowledge (i.e., between 16 and 24 months) assessed using parental report was predictive of later vocabulary, phonological awareness, reading accuracy and reading comprehension 5 years later (i.e., when children were between 4 and 9 years of age). Thus, it is possible that vocabulary in infancy is a platform for developing reading accuracy and reading comprehension skills.

To gather direct measures of children's comprehension vocabulary, researchers have used two types of measures: looking responses and touch (or haptic) responses. The LWL paradigm has been used to measure children's lexical abilities using children's looking responses after being exposed to a spoken word. This approach uses visual images to test children's knowledge of a word based on the looking patterns of the child, usually visual fixation or overall proportion of looks to target (Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Meints, Plunkett, & Harris, 1999; A. Fernald et al., 2001). The downside of this work is the labor-intensive coding of the video data as they don't use automatised tools to measure child's gaze. For this reason, the LWL paradigm has been often adapted to be automatised with the help of eye-tracking techniques. An alternative is to use touch responses to measure children's word comprehension skills. The Computerised Comprehension Task (CCT; Friend & Keplinger, 2003) is a touchscreen-based assessment that measures children's comprehension using children's touch as a response to a prompted word. A significant contribution of this task is that it is administered in an engaging portable interface with easy data extraction, facilitating data collection in children up to 20 months. The CCT has been validated and adapted to other languages and populations such as children learning Mexican Spanish (Friend & Keplinger, 2008), low and middle SES Parisian French toddlers (Scaff et al., in press), low SES and middle SES Argentinian Spanish children (Rosemberg & Alam, 2021), French-English bilingual populations (Legacy, Zesiger, Friend, & Poulin-Dubois, 2018) and even multilingual children varying in SES (Fibla et al., 2016). Using the CCT task, Friend et al. (2018) found that directly assessed vocabulary comprehension in the 2nd year of life was predictive of language skills during the 4th year of life, when children were in kindergarten in English monolingual, French monolingual and French–English bilingual children. These results support the idea that early vocabulary may provide a foundation for later vocabulary and kindergarten readiness. Some studies have started measuring both speed of word processing and online word comprehension in the same task. Smolak et al. (2021) explored if decontextualized vocabulary (measured with the CCT task as the number of correct touch responses) and speed of word processing (measured as latency to fixate the target and latency to touch) at 2 years of age predicted vocabulary during the preschool period. Results reveal that at 2 years of age, vocabulary and visual response latency (but not haptic response latency) predicted vocabulary at 3 and 4 years of age. Further, decontextualized vocabulary remained a significant predictor when controlling for speed of processing, but not vice versa. This suggests interesting relationships between vocabulary, speed of processing and later language outcomes. For instance, the number of word–referent associations and the efficiency with which these are processed are important to vocabulary outcomes, but vocabulary seems to predict later skill more accurately in these age ranges.

The CCT task has also been used to measure the influence of minimal language exposure and socioeconomic status (SES) on early word comprehension (Deanda, Arias-Trejo, Poulin-Dubois, Zesiger, & Friend, 2016). Results from this study showed that minimal second language exposure and SES exert significant and independent effects on a direct measure of vocabulary comprehension in English-dominant and English monolingual 16-month-olds. This effect was also found in a sample of Spanish-dominant and Spanish monolingual children, but there was no effect of SES on vocabulary comprehension. These results emphasize the sensitivity of the language system to minimal changes in the environment in early development, as well as the need to expand language studies to other populations where contextual variables such as SES might show different relationships to language. In fact, in western populations vocabulary size appears to be the aspect of language most sensitive to the effects of SES. Hart and Risley (1995) well known study documented differences in vocabulary size among children of professional, working class, and low SES families that increased with development. By 3 years of age, the higher SES children had produced over 1000 different words while the lower SES children had produced half that many. Literature using spontaneous speech, maternal report, and standardized tests to assess productive and receptive vocabulary have also found SES-related differences, with the size of the difference in vocabulary depending on the size of the difference in SES represented in the sample (Dollaghan et al., 1999; Hoff, 2003; Pan, Rowe, Singer, & Snow, 2005; Rowe & Goldin-Meadow, 2009). These studies emphasize the role of contextual variables on vocabulary and the need to study these relationships in more depth.

These studies provide evidence that early vocabulary is predictive of later language skills and that early vocabulary is highly influenced by contextual variables such as SES in some western populations. Moreover, direct measures of vocabulary size seem to be more predictive of later vocabulary abilities and allow one to tap into the mechanisms behind word comprehension. Finally, combining two predictive measures in a single task, such as speed of word processing and online word comprehension, allows researchers to examine how they influence each other. The present study builds on this literature to create a culturally valid online task that uses children's looking patterns, rather than touch responses, to measure both speed of word processing and word comprehension early in vocabulary development, including other early language processing skills shown to be critical in early vocabulary development. The advantage of administering these tasks using looking measures rather than touch measures is that this allows us to potentially test very young infants who might lack the skill to produce a touch after hearing a target word or that do not have the experience that most western children have with touch-screen devices. Moreover, it allows tracking of looking patterns by examining how gaze changes in response to speech over time (e.g., Mahr & Edwards, 2018), which gives fine grained measures of children's processing skills. This might be particularly informative when using language tasks across large age ranges or in populations where there is no normative work.

### 4.1.2 Other Critical Early Measures

The previous studies show that children who rapidly recognize and interpret familiar words are able to learn more words which translates into larger vocabularies. This provides indirect evidence that lexical processing efficiency is related to word-learning ability, in other words, the ability to map new words into new referents. A study by Lany (2018) found direct evidence of the relationship between lexical processing and novel word learning in 18 and 30 months children. In this study, children who were faster at recognizing familiar words were also more accurate at recognizing novel words in a word learning task. Thus, when learning a word, it is crucial to be able to quickly recognise words to build new word-object associations and build vocabulary. But learning a word is not that simple. There are additional processes that play a role. The literature shows that children tend to map novel names to novel referents rather than to familiar ones, and that prior lexical knowledge and biases towards novelty may help with this (Mather, 2013). Children's tendency to attend to a novel object when a novel word is produced in the context of both a familiar and a novel object has been explained in multiple ways. A possibility is that children use a strategy called "mutual exclusivity" based on prior lexical knowledge to determine the referent of a novel word (Markman & Wachtel, 1988). Mutual Exclusivity has been demonstrated in children from 14 to 30 months using multiple paradigms including 2-dimensional images on a screen (Bion et al., 2013) and 3-dimensional objects (Horst & Samuelson, 2008). The use of mutual exclusivity to determine a referent has been shown to be driven by how well the child knows the familiar objects presented with the novel object. Children are able to disambiguate between a familiar object and a novel one when presented with a novel word if they have a strong association between the familiar object and the word that defines it. That is, they are able to map a new word into a novel object when that appears in the context of a highly familiar object. Thus, in mutual exclusivity, children bring their previous knowledge to bear in-the-moment to select the referent of a novel word (Samuelson et al., 2017).

Children's disambiguation skills have also been explained by their more general attraction to novelty; a phenomenon known as "novelty bias". In western cultures, the interaction between novelty driven attention and lexical knowledge has been evaluated in referent selection and retention tasks (RSR). Studies using the RSR task show the complexity of the role of novelty bias in word learning. This is mainly due to two reasons. First, studies show that children's attention to novelty continuously changes over development. Second, even though attraction to novelty might be very useful when learning a new word-object mapping during referent selection, too much attraction to novelty could prevent retaining that new object-word association. Kucker et al. (2018) found negative associations between attention to novelty and retention of new word-referent links across individual 18-months-old children using 3D objects using the RSR task. This study also examined possible sources of bias though a computational approach, their results suggests that when lexical knowledge is weak, attention to novelty does not help learn the new word (i.e., retention of the novel-object word association). In another study using a very similar version of the RSR task, 24 month-old children overcame the novelty bias and correctly selected a novel referent in response to a novel word, but they could not remember it after a 5 minute delay (Horst & Samuelson, 2008; Bion et al., 2013; Kucker & Samuelson, 2012). By 30-months of age, however, children were able to overcome the novelty bias to select the correct referent and remember novel name–referent mappings over a time delay (Bion et al., 2013; Spiegel & Halberda, 2011).These studies suggest strong age effects on the ability to remember word-object associations.

Retention of new words however, is also affected by the strength of children's lexical representations (Kucker & Samuelson, 2012). In this study, a short pre-familiarisation was enough to boost retention of the novel name-referent mappings formed during referent selection at 24 month-old but 18 month-old did not show retention. Moreover, Bion et al. (2013) found a significant correlation between CDI productive vocabulary and disambiguation skills (even when controlling for age). This suggests that referent selection and retention might be related to previous learning experiences. A recent study also supported this idea; weaker vocabulary knowledge during the initial exposure to a new word led to better retention of new mappings (Kucker, McMurray, & Samuelson, 2020). In another study, referent selection performance was significantly reduced on trials with weakly known competitors. However, children showed above-chance retention for novel words mapped in the context of weakly known competitors compared with those mapped with strongly known competitors or with completely novel competitors. This highlights the relevance of the strength of known lexical representations relative to attraction to novelty, highlighting the importance of accurately measuring vocabulary knowledge early in development (Samuelson et al., 2017; Kucker et al., 2018).

The studies reviewed above show the importance of lexical knowledge in relation to novelty biases, referent selection and retention of new words. However, the role of novelty in word learning has never been explored in other cultures. It is possible that non-western cultures show different patterns of attraction to novelty across early development. Examining this relationships in children growing up in other cultures could help better understand the relationships between novelty biases and lexical knowledge in word learning. Moreover, in the western context vocabulary has proved to be highly sensitive to contextual variables such as SES, we could imagine that children from lower SES backgrounds, with reduced language learning opportunities, have less instances to put into practice referent selection abilities and therefore show less skills in RSR. The reviewed literature used participants from middle or high SES, but there is one study has examined the association between vocabulary knowledge and fast mapping skill in low-SES preschoolers. Fast mapping refers to the mental process whereby a new concept is learned based only on minimal exposure to a given unit of information (Carey & Bartlett, 1978), such as in RSR tasks. This study did not find a significant correlation between PPVT scores and performance on a fast mapping task (E. J. Spencer & Schuele, 2012). It is possible that the link between language learning process skills and vocabulary knowledge is weak in low-SES children, maybe due to limited language exposure. Additionally, this study examined fast mapping of terms for object parts and not whole objects as the literature previously reviewed. Thus, it is not clear the role of SES status in RSR abilities, neither what developmental patterns should be expected in very different cultural settings.

Altogether, the previous studies show that children need to master a set of skills to be able to effectively learn a word. Evidence also suggests relationships between speed of word processing and vocabulary abilities, as well as speed of word processing in relation to word learning abilities. Combining these measures in a single task such the ELP allows researchers to examine how they influence each other. This multi-factor view of early word learning could explain the large individual variation we see in early language development (e.g., Frank et al., 2017). Moreover, some of these language processes are affected by children's environment, in particular SES in western samples. This means that ELP might be able to capture individual differences based on environmental variables such as SES. The challenge is that to date most data on early word learning comes from western, typically-developing samples. There are relatively few studies examining non-western samples and, even fewer, with samples growing up in poverty. The present study aims to extend this literature beyond these western contexts to look at a high-risk population of children learning the Awadhi language in rural India.

### 4.1.3 The present study

The specific goals of this study are: 1) To adapt the ELP task to a new language and culture; 2) To use the ELP task to gather measures of language processing in Indian children at different ages; 3) To measure language processing abilities at 30 months of age in the same sample.

After developing the ELP task for use with children living in the UK learning British English, we adapted the task to Awadhi and the cultural and social context of our sample in rural India together with our collaborators from the Community Empowerment Lab (CEL), a local organisation. We wanted the ELP task to be culturally relevant to our target population and thus, instead of a mere translation of the task (from English to Awadhi), we aimed to create an adaptation of ELP that would be suitable for testing language processing abilities of children growing up in Shivgarh, India. Thus, we adapted the task using an iterative process with our local collaborators. Also, we wanted the task to be equally valid for using it in the UK, with the only difference being the language of the task (British English or Awadhi); thus, we carefully selected words and images that would be appropriate across both sites.

While generalising the ELP task to the Indian population, we faced several challenges. The first one was what words to include in the task. As in the ELP version developed in the UK, we aimed to have highly familiar words as well as a range of difficulties (i.e., easy, moderate and difficult words), which would allow us to capture individual differences in our sample. However, there are no normative measures nor normative data on children growing up in rural India learning Awadhi. Thus, we did not have a reference dataset to establish a baseline distribution for language scores or language measurements including word frequencies. Consequently, we started by using the words we had for ELP in British English and assessed how relevant they were for our purposes. Together with the CEL team, we gathered measures of word frequency via adult (mothers) checklists. That gave us an estimate of how frequent our selection of words were. Highly infrequent words were excluded.

Next, we translated the carrier phrase used in ELP UK into Awadhi and adjusted both forms to make them as similar as possible making sure the target word appeared at the end of the sentence in the following way: "Look, where is the (target)?" in English and "Deko, kahan hai (target)?" in Awadhi. We recorded both carrier sentence and words with a native female speaker of Awadhi. To make sure our image selection was culturally relevant, the CEL team members rated the images. We excluded all images that were not appropriate and changed them for culturally relevant ones. Once we had a selection of words and images that seemed appropriate, we used adult report (mimicking the ELP task) to test the relevance of the audio recordings and images. That is, we showed the final selection of image pairs to adults from Shivgarh, and played the audio recordings of one of the words. We asked them to point to the correct image and give us feedback. Instances in which adults found the word and the image correspondence ambiguous or difficult, were revisited or excluded. After we selected the most appropriate images and words and added them to the eye-tracking task, we piloted the ELP adaptation with children between 18 months and 6 years. We used these data as well as our observations to inform our final selection. Once the final version of ELP in Awadhi was created, we revisited the UK version so both tests would include the same images and words in each language respectively (see https://osf.io/yczgj/ for the final selection of images, words and audio recordings of ELP in British English and Awadhi).

After adapting ELP to Awadhi, children's language processing abilities were assessed. We chose to test a large age range (from 17 to 48 months of age), because previous data from CEL using an adapted version of the ASQ®-3 (The Ages & Stages Questionnaires, Third Edition; Squires et al., 2009) suggested a possible delay in langauge abilities. However, our collaborators highlight that many aspects of the ASQ®-3 adaptation were not culturally relevant in that population. To be able to capture developmental trends across individual children in our sample, some of the younger children were tested again when they were approximately 30 months of age. Critically, the ELP task is suitable across a large age range, and it was designed to be a portable task, which allowed us to pack the eye-tracker in a backpack, fly with it to India, and test in a rural location.

## 4.2 Methods

### 4.2.1 Participants

The final sample for the 18-month-old group included 177 children aged 17- to 48-months-old (*M* = 24.29 months, *SD* = 8 months, 78 female, 94 Low SES). Table 4.1 shows the sample demographics. An additional 64 children were recruited to participate in the study but were not included in final analysis due to fussiness (10), canceled the session (5), technical problem (2), did not provide SES information (28), did not provide enough usable data (e.g., had noisy eye tracking data, 19). This means that we obtained usable eye-tracking data on 205 out of 241 children we tested, which is 85.06%. This is impressive given the rural setting in which we collected this data, with children who have rarely seen a TV screen and are not used to sit still in a chair. The final sample for the 30-months-old retest group included 41 children between 34 and 37 months of age *M* = 35.57 months, *SD* = 0.73 months, 28 female, 21 from Low SES). An additional 28 children were recruited to participate in the study but were not included in final analysis due to fussiness (2), canceled the session (3), did not provide have SES information (1), did not provide enough usable data (e.g., had noisy eye tracking data, 22). A subset of 27 children had usable data at both time points (16 female). All participants were born full term. None of the participants or their mothers had been diagnosed with any major psychiatric illnesses or had unusual characteristics as observed by the CEL research staff.

This work was supported by Grant No.OPP1164153 from the Bill & Melinda Gates Foundation. The project was reviewed and approved by CEL ethics committee. Parents signed or provided an oral videotaped informed consent form (in cases where the caregiver was illiterate). The subset of the data re-

| **Sample Demographics**<br>overall N = 177 (78 girls) | |
|---|---|
| **Age in Months** | |
| Mean (SD) | 23.2 (7.4) |
| Median [Min, Max] | 21 [17, 48] |
| **SES** | |
| Low<br>(illiterate or primary education) | 94 (53.1%) |
| High<br>(greater than middle school) | 83 (46.9%) |
| **Caste** | |
| Scheduled caste-scheduled tribe<br>(traditionally most depressed) | 93 (52.5%) |
| Other backwards caste<br>(socially or economically disadvantaged) | 39 (22.0%) |
| General<br>(middle class) | 8 (4.5%) |
| not specified | 37 (20.9%) |

**Table 4.1:** Summary of sample demographics for ELP

ported here is part of a larger study examining the early precursors of executive function led by Prof. John Spencer. The 2018 Patrice L. Engle Dissertation grant to Laia Fibla Reixachs provided support for carrying on this research in India.

### 4.2.2 Procedure

The ELP task was carried out on a 24-inch BenQ Zowie XL2411P monitor screen that was connected to a Gigabyte mini computer used to display the stimuli and a Lenovo laptop host that interfaced with the eye-tracker software running SR Research Experiment Builder. Participants were seated on their caregiver's lap on a sofa approximately 80 cm from screen. The eye to camera distance was about 50 cm, and the eyes were in line with the top part of the screen. The eye-tracker was positioned in the horizontal center of the screen. The eye tracker is an Eye-Link Portable Duo (SR Research, Ontario, Canada) in the remote setting. Both screen and eye tracker were placed on a table (together with the Gigabyte mini computer and the Lenovo). As part of the set-up there were two additional cameras in the room, one located on top of the monitor using a tripod, which recorded the participant's face (a GoPro model HERO5) and one located at the back of the room to record the experiment as it was presented on the monitor. These recordings were done to monitor and keep a record of the participant doing the task. Our portable setup also included a foldable silicon keyboard, a mini Xmi Pte Ltd portable speaker and a standard computer mouse. The setup is shown in Figure 4.1. A small target sticker was placed on participants' foreheads which allowed tracking of head (and eye) position even when participants moved or the pupil image was lost. The eye tracker was set to monocular recording such that it tracked the gaze position of a single eye using pupil and corneal reflections of an infrared light source. The sampling rate was 500 Hz.

Due to the portable aspect of this setup, we also allowed the experimenter and the laptop that monitored the experiment to be in the same room. We trained the experimenter to not distract, give feedback or engage with the participant during the task. All participants were tested in the company of a community member who was part of the CEL team. This team member was either trained as an experimenter or trained to accompany the mother-child dyad during the task. They also helped provide translations or explanations related to the task to the participants.

The experiment began with a short clip of *Fantasia, 1995* (Disney). While this video played, the experimenter placed the small target sticker on the participants' forehead. Once the target sticker was in place, the tracking camera


**Figure 4.1:** Portable ELP setup in India: 1) participant 2) eye tracker 3) screen 4) participant view camera 5) setup view camera 6) computer interfacing eye tracking software 7) computer running experiment.

was adjusted so the distance from target to camera was approximately 50 cm. The experimenter adjusted the participant's position as many times as needed so they would be placed in the most optimal position and distance. After checking that the pupil and corneal reflection were visible on the camera, the calibration procedure began. During calibration, participants were shown a looming black and white geometric shape in five locations of the screen (middle, top, bottom, left, right) used to map raw eye position data to the camera image data. This allowed mapping of gaze position to the stimulus presentation. Following successful calibration, the experiment commenced. The ELP task was divided into two blocks separated by a 5-minute break. The calibration procedure took place twice during ELP - at the beginning of the task and after the 5-minute break.

Because the ELP task is an image choice-based task, each trial started displaying two pictures on the screen for 2000 ms. Then, the screen was covered by a gray transparent filter and a gaze contingent cartoon appeared in the center. When the child looked at her, she named the target embedded in a carrier sentence such as ”Deko, kahan hai (target)?” (”Look, were is the (target)?”). At the onset of the target word, the gray filter disappeared and the child could clearly look at either the named image (target) or at the other one (distractor). The pictures remained for a 3200 ms response period and finally there was a reward which consisted of the cartoon happily jumping up and down. This positive reward was always displayed (see the left panel of Figure 4.2 for the general structure of the ELP trial). In the ELP task, all image pairs were matched in salience and complexity.

The ELP task includes five measures of language processing: speed of word processing, word comprehension, novelty bias, referent selection and retention of new words. Those measures were incorporated in the task using four different types of trials (see examples of each trial on 4.2):

- **Reinforcement or Speed of Processing Trials:** Contain two pairs of highly familiar nouns (*flower-ball* and *baby-dog*) that repeat 5 times each during the task. Two different sets of images were used in this trials to keep children interested in the task.
- **Comprehension Trials:** Include 41 pairs of nouns, verbs and adjectives varying in difficulty.
- **Referent Selection Trials:** Contain 8 word-image pairs with one well-known and one novel object. The novel is the target on 4 of those trials, and the familiar on the other 4. Half of the familiar images used in Referent Selection trials were extracted from Retention trials (i.e., they were highly familiar nouns to which the child was exposed five times before

seeing them in the context of a novel image). The other half were familiar nouns that also appeared in Comprehension trials, the child saw them only once before they appeared paired with a novel). This allowed us to manipulate image familiarity, highly familiar (images from Reinforcement trials) versus familiar (images from Comprehension trials).

- **Retention Trials:** Included images pairs of two previously-mapped novel objects. In the 4 retention trials, the child was asked to look to the novel word-image they saw during Referent Selection trials. Children were also exposed to the distractor novel image because it appeared as foil on Referent Selection trials where the target was the familiar.

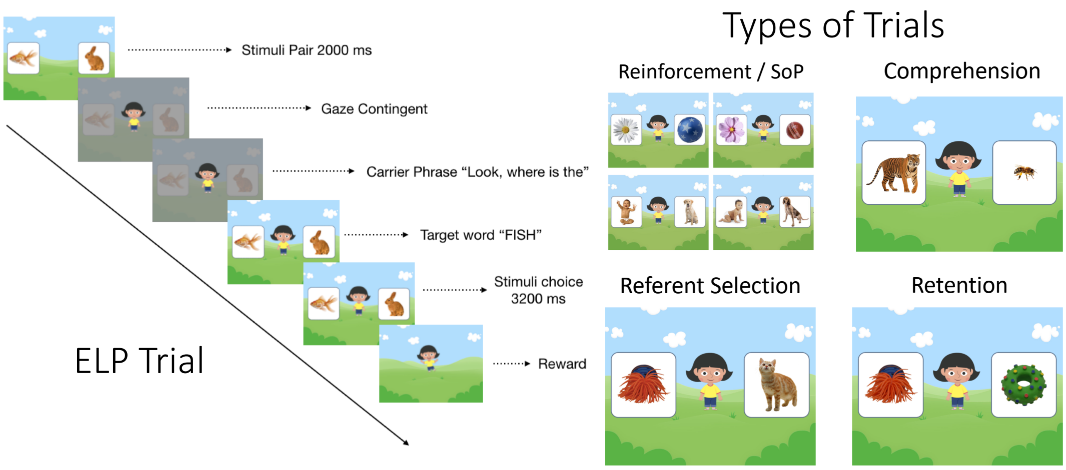


**Figure 4.2:** Trial schematic for the ELP task including trial types with examples.

The ELP task consists of two blocks separated by a 5-minute break or retention interval during which children can either stand up or watch a short movie on the screen (Piper a 2016 computer-animated short film produced by Pixar Animation Studios). Children started ELP with a the first block that contained 5 reinforcement trials mixed into 20 comprehension trials followed by 8 reference selection trials (4 in which the target was the novel, and 4 in which it was the familiar). The different types of trials in block 1 were presented in the following order: 2 reinforcement trials, 7 comprehension trials containing easy and moderate nouns verbs and adjectives, 1 reinforcement trial, 8 comprehension trials with easy and moderate nouns verbs and adjectives, 1 reinforcement, 5 comprehension trials with easy and moderate nouns verbs and adjectives, 1 reinforcement, 8 referent selection. After the 5-minute break and the second calibration (in case the child stood up during the break or moved from the initial position), the second block started. Children then were exposed to 5 reinforcement trials mixed with 4 retention trials and 20 comprehension trials. Trials in block 2 were presented in the following order: 2 reinforcement trials, 4 retention trials, 1 reinforcement trial, 5 comprehension trials containing moderate and difficult nouns, verbs and adjectives, 1 reinforcement trial, 8 comprehension trials with moderate and difficult nouns, verbs and adjectives, 1 reinforcement trial, 7 comprehension trials with moderate and difficult nouns, verbs and adjectives. For Comprehension trials, the first block only contained easy and moderate words, whereas the second block only contained moderate and difficult words. This meant that the task increased in difficulty as the child went through it. This was designed in this way to be able to test children from a large age range and still get enough data from younger and older children.

In each block, word order was pseudo-randomised to ensure that the target did not appear on the same side of the screen more than two trials in a row, and that the word type/difficulty did not repeat more than two trials in a row. Referent selection and Retention trials were randomised separately but followed the same criteria, that the same word type would not appear more than two times as the target (for Referent Selection trials), and that the target would not appear on the same side more than two times (for both Referent Selection and Retention trials). Thus, we had two fixed pseudo-randomised

ELP versions named order 1 and order 2. To keep the task short, for each image pair, children were only asked for one of the images (but not the other). This led to creating two different target word versions A and B. For example, in the word pair *cat* - *fish*, order A asked the child to look at *cat*, and order B asked the child to look at *fish* (but in order A *fish* was never the target and in order B *cat* was never the target). This meant that the ELP task had four different versions based on target word and randomisation: A1, A2, B1 and B2. We tested approximately the same amount of children in each order and checked for possible order effects in our analyses. Further details regarding the ELP setup, the task and how we developed it can be found in the Methods section of Chapter 2.

#### 4.2.2.1 Adapting the ELP task to the Indian Context

We adapted the UK version of the ELP task to the Indian context and local language. Words and carrier sentences were translated to Awadhi. When translating the carrier phrases, we chose the form that was most equivalent across languages (English and Awadhi), because we aimed to create two ELP versions that were as similar as possible across sites. This meant some adaptations in the original English version. For example, in the initial version, the ELP noun carrier phrase in English was ”Look at the (target)!”, however this did not work very well in Awadhi and thus, we modified it to be ”Look, where is the (target)?” in English which in Awadhi would be ”Dyakho, kahan hai (target)?”, which has a quite similar form in both languages. For verb sentences, we used ”Look, who is (target)?” in English and ”Dyakho, ko aay (target) raha hai?” in Awadhi. For adjectives, we used ”Look, which one is (target)?” in English and ”Dyakho, ka/kon (target) hai” in Awadhi (for the interest of the reader we provide the equivalent English forms of these sentences split by word: *Dyakho* (Look), *kahan* (where), *hai* (is), *raha hai* (-ing), *ka/kon* (what/who), *ko aay* (which one)).

The audio stimuli were recorded using a female native speaker of Awadhi who was also a member of that community. We asked her to speak as she would do to a child while recordings the sentences and words. Audio recordings were recorded using the GarageBand application by Apple with a mac OS, which includes a function to remove background noise. Stimuli were later extracted from the recordings and processed using Praat scripts (Boersma, 2001). We added silences at the edges of each sound file (0.01 s on each edge), and we normalised the sound intensity (i.e., amplitude). We recorded several examples of each word with its carrier sentence. Per each recording we extracted the total duration, root-mean-square pressure (i.e., the square root of the average of the square of the pressure of the sound signal over a given duration), the intensity in decibels and the average, minimum and maximum pitch. Those measures helped inform our selection of the best and clearest examples of each of the words we recorded.

When adapting ELP to Awadhi, we also gathered measures of word frequency, based on adult report, to have an estimate of how likely the words we selected for ELP were to be used in children's daily environments. Together with our local collaborators, we created a set of questionnaires that mimicked a vocabulary checklist in which we included the selected ELP words. Then, we asked mothers from Shivgarh to rate if the words were usual in their child's daily life by answering if their child heard, understood, understood and said, or didn't hear each of the words. Highly infrequent words were excluded from the ELP task and replaced by words that children were exposed to. We allowed for a large frequency range so we could have some highly frequent words, considered very easy for children, and some harder words.

Images were validated by several members of the local team in India, to assess the relevance for the cultural context. Inappropriate images were replaced. Because we wanted to use the same ELP task across site, images were selected to be culturally relevant in both UK and India. Cases in which that was not possible were excluded from the task. For example, *bed* and *tap* were excluded from ELP because they look extremely different across sites. They were replaced by *door* and *basket* using images that could be identified in both cultures.

Once a first Awadhi version of ELP was created, we piloted it with adult members of the community as well as with older children. This allowed us to have feedback on the different aspects of the task. Everything that was not culturally relevant or appropriate was changed or removed. Once the Indian ELP version was ready, we revisited the UK version, so both tests included the same images and words, and well as a carrier phrase that followed the most similar structure possible (the final stimuli selection can be accessed in https://osf.io/yczgj/).

### 4.2.3 Analysis Method

The eye-tracking data were pre-processed using Data Viewer (SR-Research, Ontario, Canada). Trials were segmented into periods of interest (IP) using message-based events. Areas of interest (AOI) were set to be 50% bigger than target objects to account for calibration errors and drifts in the eye tracker. Sample reports were exported and raw gaze position was further analyzed using the statistical package R (R Core Team, 2017), as well as eyetrackingR (Dink & Ferguson, 2016), an R package designed to work with eye-tracking data. A common measure in eyetracking studies of word recognition is an accuracy growth curve (also called Growth curve analysis (GCA); Mirman, 2014). The growth curve measures how the probability of fixating the target changes over time. We computed this growth curve using eyetrackingR. Looking to the target side and the distractor at each point in time during the trial was aggregated into 100ms time bins allowing calculation of the proportion of looks to the target. We ignored off-screen looks or looks out of our AOI when computing this proportion. Trials were only included in the analysis if they had more the 60% of looking data.

Out of the total ELP trial, our analyses focus on two windows of interest: one during the familiarisation phase, and the other during the test phase. We chose these two windows of interest with two objectives. First, we wanted to measure if children had a preference for any of the images before hearing the target word, that is, during familiarisation. Second, we wanted to measure if children looked at the target image after hearing the target word during the test phase. Looking data during familiarisation included looks towards the two images before hearing the target. To allow for the best possible statistical modelling of these time series data, the looking data from the first 300ms of the familiarisation phase of the trial was trimmed to reduce noise. Looking data from the test phase focused on a window of interest that went from word onset to 1800ms, consistent with previous studies. This criteria is based on previous literature suggesting that 24-months-old children shift their gaze to the correct picture before the end of the spoken word, in contrast to 15-months-old who do not orient to the correct picture until after the end of the target word (e.g., A. Fernald et al., 2008, 2001). Since we had a large age range, we wanted to take into account looks to target from word onset rather than word offset, since that would capture age effects in processing abilities.

We had four specific questions we wanted to address: Q1) how did children perform as a group in the different ELP measures at Test (i.e., when they were approximately 18 months)?; Q2) how individual children perform across ELP measures at Test (approximately 18 months)?; Q3) how ELP performance at Test relates to performance at Retest (at approximately 30 months) at the group level?; and Q4) how ELP performance at Test relates to performance at Retest the individual level?

To answer Q1, proportion of looks to the target through time were fit with a binomial hierarchical model estimated with a Laplace approximation using the glmmTMB package (Brooks Mollie et al., 2017) and eyetrackingR (Dink & Ferguson, 2016) in the statistical package R (R Core Team, 2017). The model was fit with orthogonal polynomials of the time term following the growth curve analysis approach (GCA) (Mirman, 2014), that is, the data were modelled with Time, time squared, up to time to the power 4, but scaled and centred so as to not be correlated with one another. In addition, the model contained fixed effects of Age in months represented as a continuous variable, Trial Type which included all ELP types of trials, and SES based on maternal education (low = illiterate or primary education; high = greater than middle school). SES was scaled and centered. Each of the time terms were nested as a random effect within participant, along with allowing each participant a random intercept for a maximally-specified model. The model was fit with Age, Gender, Trial Type and SES as predictors. The model was then simplified using the Akaike information criterion (AIC), an estimator of prediction error, and the Anova function of the R package (Wagenmakers & Farrell, 2004). Because Gender did not show any consistent results nor improved the AIC of the model, that is, Anova comparisons between a model that included gender and a model that did not were not significant, it was removed from the models. Models were also tested using the DHARMa R package (Hartig, 2021), which creates readily interpretable scaled (quantile) residuals for linear mixed effects models, as well as plot and test functions for typical model miss-specification problems (e.g., over/underdispersion, zero-inflation, and residual spatial and temporal autocorrelation). The same model was used to answer Q3 but, because in this case we modelled ELP data at both time points (ELP test at 18 and ELP retest at 30 months), the model also included test type (test, retest) as fixed effect as well as a part of the random effects structure (interacting with participant). This big initial model was used to assess overall differences between ELP measures in a single model. The same model (with the same fixed effects and random structure) was used to look at proportion of looks to the target through time, split by each ELP trial (or measure) at both 18 and 30 months, as well as to control for image preference during familiarisation across the different ELP trials. Variations from this model are detailed in the pertinent section in Results. Any variation from this model aimed to better fit the data. Changes from the original model were assessed using the AIC criterion and Anova comparisons (Wagenmakers & Farrell, 2004), as well as DHARMa plot and test functions (Hartig, 2021), for typical model miss-specification problems using R package. These were the same methods applied to evaluate the best initial model.

To answer Q2 and Q4, we used correlation analyses. For that, we computed the overall mean proportion of looks to target on ELP test during the window of interest (from word onset to 1800 ms). We used eyetrackingR (Dink & Ferguson, 2016) to compute this proportion. Q2 focused on individual performance across ELP measures and, thus, we ran a set of correlations between the different ELP trials at 18 months. To correct for multiple comparisons, we set a more conservative criteria and only considered effects with a significance level smaller than 0.01 (sig.level <0.01). Q4 looked at individual performance across both ELP observations and, thus, we run a set of correlation analysis to measure relationships between ELP at 18 months (test) versus ELP at 30 months (retest). Correlations were run in the R package (R Core Team, 2017).

## 4.3 Results

Our participants did ELP both at Test and Retest (approximately 18 and 30 months of age). In this section, we examine the relations between the different ELP measures at both time points. We first present results on the ELP data collected at Test (18 months), followed by results on ELP data at both Test and Retest (18 and 30 months). We use both GCA approach and correlation analyses to measure performance at the group level and at the individual level at 18 months, and performance at the group level and at the individual level across both time points (18 and 30 months).

### 4.3.1 ELP Test at 18 months

Here we report results in answer to Q1: how did children perform as a group in the different ELP measures at Test? We then examine Q2: how did individual children perform across ELP measures at 18 months?

Looking proportions to the target across ELP trials were modelled following the GCA approach. The model was fit with Time in trial using orthogonal septic polynomials of the time term up to the fourth order (ot1, ot2, ot3 and ot4), Trial Type, Age (in months) and SES as fixed effects. Time and Participants were also added as random effects. The aim was to assess if the task was sensitive to differences in performance across the language processing measures included in ELP, while taking into account age and contextual variables such as SES.

Results showed main effects of the linear, quadratic and quartic Time terms, Age, Trial Type and SES as well as multiple interactions among the variables. Here, we focus on the significant 4-way interactions between the linear, quadratic, cubic and quartic Time terms and Age, Trial Type and SES. A full report of the results can be seen in Table 4.2 at the end of this chapter.

These results indicate that children's rate of looks to target (linear time term) increase over the ELP trial as a function of Age and SES, as well as the Type of Trial that children were performing. As can be seen in Figure 4.3, younger children were mostly at chance levels for all ELP trial types. However, older children, particularly those from higher SES families, looked significantly more to target over time. Also, we can see that older children looked more quickly and more robustly to target in reinforcement and referent selection trials, a bit less in comprehension trials, and even less in retention trials. Again, this pattern was more pronounced in older children from high SES families. These results indicate that the Indian adaptation of the ELP task is sensitive to age effects as well as contextual variables such as SES, and it is able to capture differences across trials (or ELP measures) in our Indian population at older ages.

Since the overall model showed evidence of differences across ELP trial types, we modeled each trial type separately. We used the GCA approach splitting by trial type with some modifications as noted in the sections below.

### ELP Reinforcement

Looking proportions to the target for ELP Reinforcement trials both during familiarisation and test were modelled following the GCA approach. We used one model where the predicted variable was proportion looking to target during familiarisation (i.e., the image that will become the target during test), and another model where the predicted variable was proportion looking to

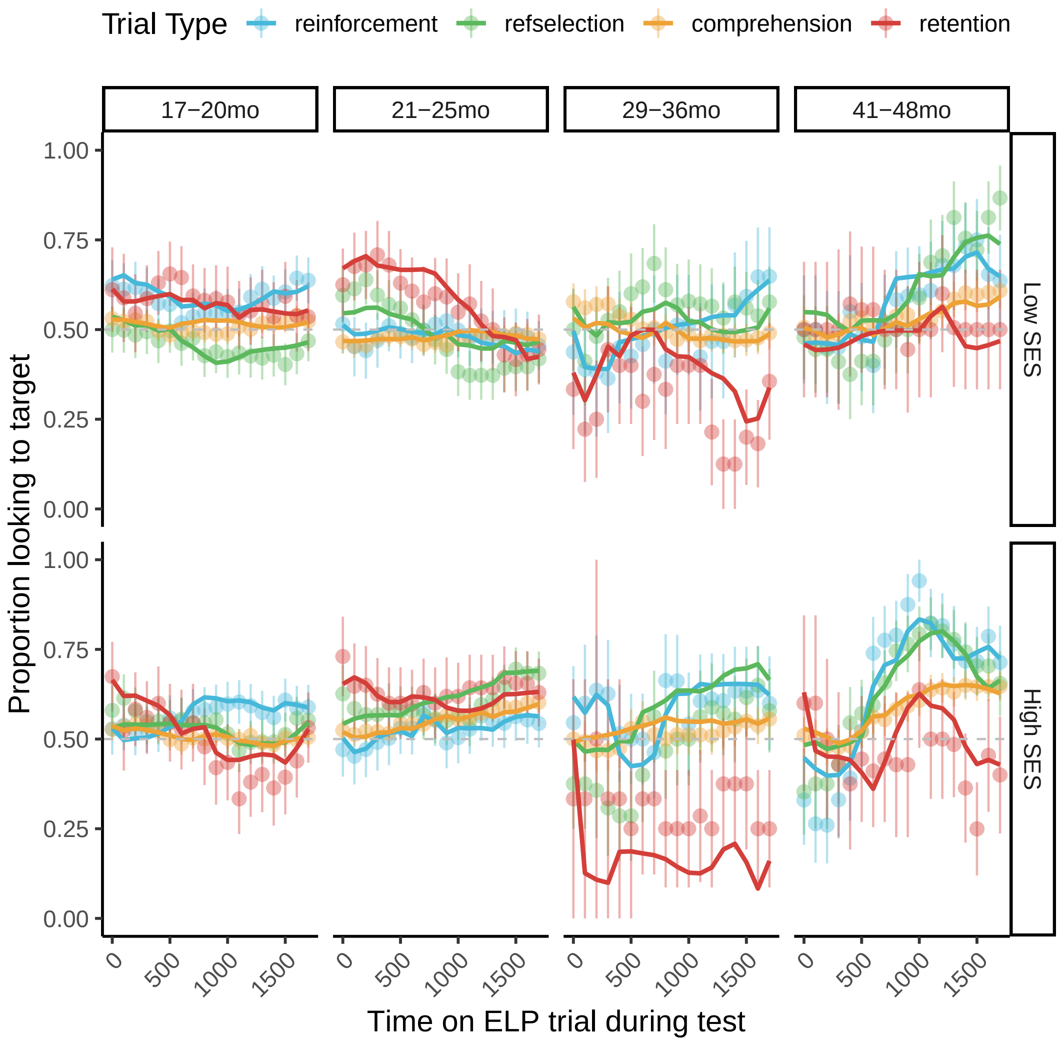


**Figure 4.3:** Model predicted proportion looking to target by Trial Type by Age and SES. Grey dashed line depicts chance performance (0.50). Age in months is split in age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

target during test. Other than the predicted variable, both models had the same structure. The models were fit with Time in trial using orthogonal septic polynomials of the time term up to the third order (ot1, ot2 and ot3), Repetition Pair Count which indicated the repetition number of the image pair, Age (in months) and SES as fixed effects. Time and Reinforcement Pair count were nested as random effects within participants, along with allowing each participant a random intercept for a maximally-specified model.

Results from the model using looking data extracted from the familiarisation phase showed a main effect of the linear Time term, some 2 and 3-way interactions between the linear Time term and Repetition Pair Count, Age and/or SES, and a 4-way interaction between the linear and the quadratic Time term, Age, Repetition Pair Count and SES. Full results are shown in Table 4.3 at the end of this chapter. As can be seen in Figure 4.4, younger children showed chance levels of looking, indicating no preference for either image, but older children show a negative trend. This indicated that they look more towards the image that will be the "future" distractor late in the trial. This negative looking pattern is modulated by children's age, their SES status and the amount of times they have been exposed to the image pair. Variation in younger ages oscillates around chance, which indicates no clear biases in looking during familiarisation. At older ages, children could be experiencing a type of priming effect, because it seems that when a new Reinforcement trial started they were looking at the image that was previously assigned as the target (i.e., the one that was last named). We also see, however, a significant 4-way interaction with the quadratic Time term which indicates that looking oscillated between images.

Results from the model using looking data extracted from the Reinforcement test phase showed a main effect of the linear, quadratic and cubic Time terms, as well as a main effect of Repetition Pair Count and SES. There are several 2-way interactions including the linear, quadratic and cubic Time terms and Repetition Pair, as well as with SES (but not between Time and Age). We also see significant 3-way interactions between Time, Age and Repetition Pair Count as well as with SES. Finally, we see a significant 4-way interactions between the linear, quadratic and cubic Time terms, Age, Repetition Pair Count and SES. As can be seen in Figure 4.5, younger children do not show a pref-

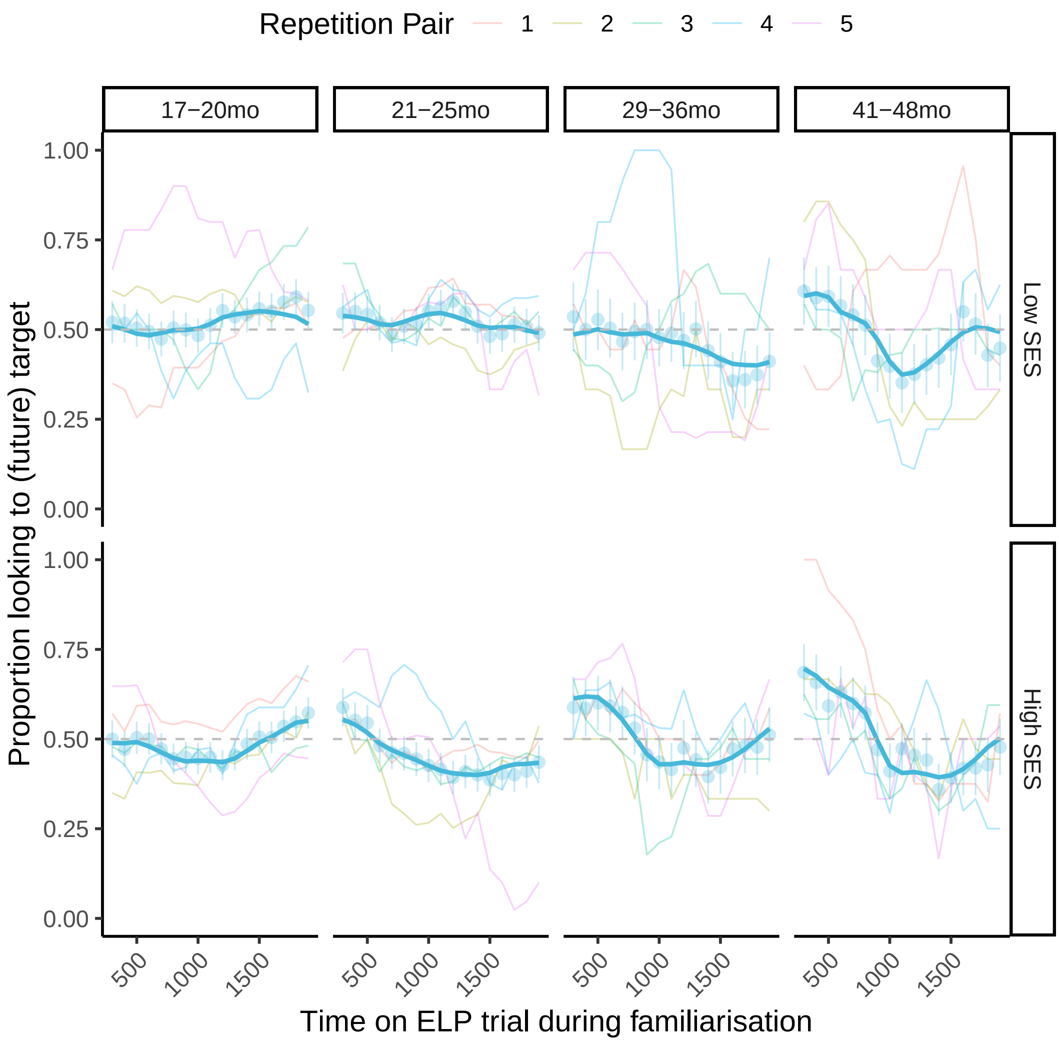


**Figure 4.4:** Model predicted proportion looking to target on Reinforcement trials by Age and SES including Repetition Pair during familiarisation. Grey dashed line depicts chance performance (0.50). Age in months is split in age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

erence for the target, however, older children show more and quicker looks to the target. This suggests that older children know the Reinforcement words better than younger children which is not surprising as they were chosen to be highly familiar. At older ages, we see an initial preference for the distractor, which reverses towards the target during the trial. This pattern might be related to our findings during Familiarisation, with children first looking at the previously named image and then to the target. This pattern seems more pronounced for low SES children as opposed to high SES children who look more quickly and more consistently to target.

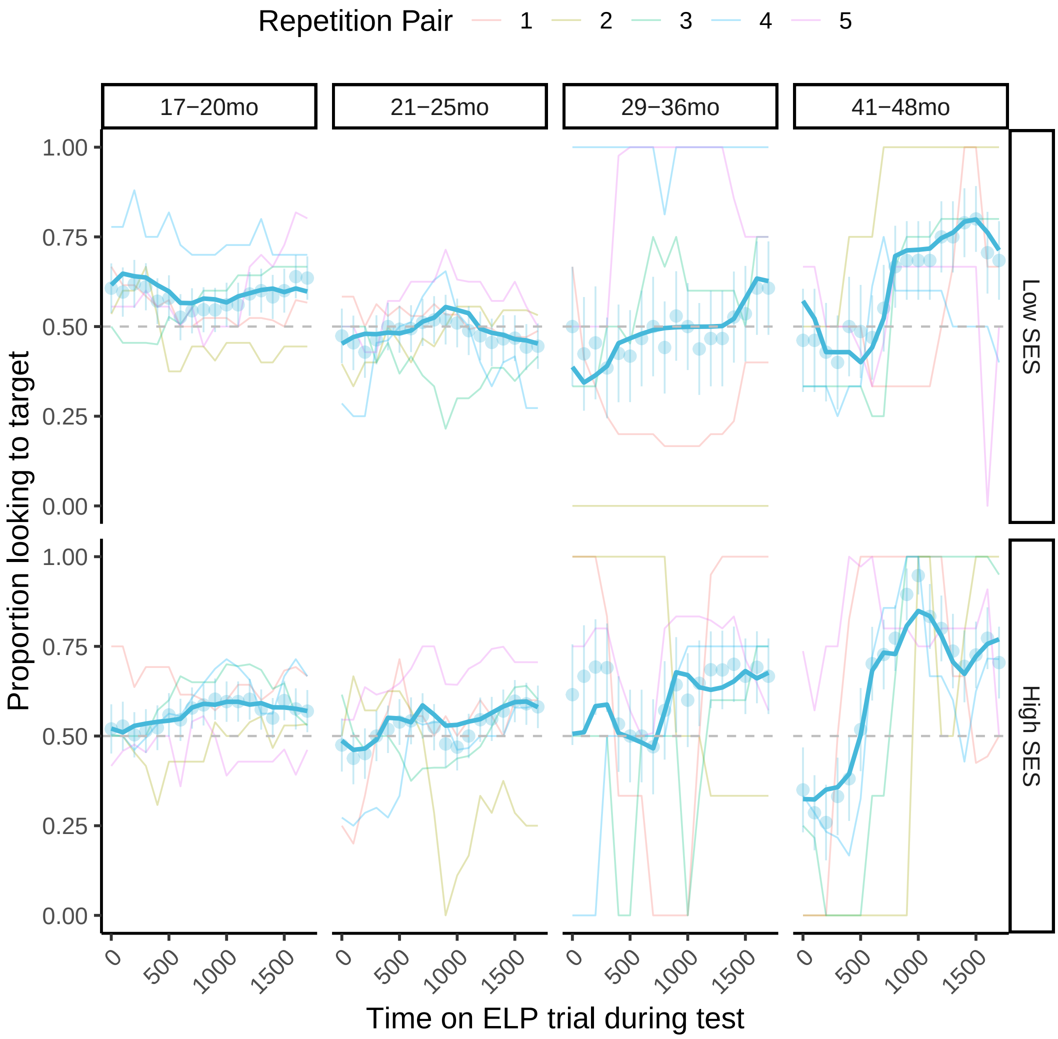


**Figure 4.5:** Model predicted proportion looking to target on Reinforcement trials split by Age and SES and including Repetition Pair during test. Grey dashed line depicts chance performance (0.50). Age in months was split in three age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

The Reinforcement trials are based on the speed of word processing measure developed by A. Fernald et al. (1998). Speed of word processing has been calculated using different scores across studies, and have often been reported to be related with each other. In the literature, *accuracy* measures of word processing efficiency are usually extracted as the proportion of looking time to the target picture during the window of interest. *Reaction time (RT)* or speed of response to the spoken word is usually calculated as the mean latency (in milliseconds) to shift from the distractor to the target picture on all distractor-initial trials on which a correct shift occurred within the 0–1,800-ms window (or 300–1,800-ms window) from target word onset (e.g., A. Fernald et al., 2006). Those two measures are usually related because children with faster shifts to the target usually also show higher accuracies. Recent studies have added new ways to calculate speed of word processing using a growth curve. In GCA, the *linear time term* captures the overall steepness of the growth curve, which can be used to quantify the lexical processing efficiency of a participant. This linear term can be extracted from model coefficients across individuals, and has been correlated with accuracy. Children who have steeper growth curves, usually also show higher accuracy in speed of word processing tasks (Mahr & Edwards, 2018).

In this study, we wanted to assess if different measures of speed of word processing related to one another in our task. Toward that aim, and following previous studies, we used looking data in ELP Reinforcement trials to extract, per each participant, the following measures: a) overall mean proportion looking to target in the window of interest (i.e., accuracy), b) the linear time term from our model coefficients on the window of interest, c) RT of first look to target. We explored these three indexes because they measure different aspects of word processing.

Our correlation analysis showed a negative relationship between ELP Reinforcement and the model coefficient (the linear time term on Reinforcement trials; $t = -6.760$; $r = -0.5129$; $p = <.000$). This relationship is not very meaningful in our data because most of our participants perform at chance in this sample (and our correlation analyses collapse across ages). Proportions of looks to target are very low and we could image that some children started the trial looking at the target and for a while stayed there (so they have a proportion higher than 0.5), but by the end of the trial, they started looking away from the target so they have a negative slope (see Appendix, Figure A.11 for more detailed values). We did not find either a significant relationship between proportion looking to the target and RT ($t = 0.760$; $r = 0.067$; $p = 0.448$).

#### ELP Comprehension

Looking proportions to the target for ELP Comprehension trials (including all word types and difficulties) both during familiarisation and test, were modelled following the GCA approach.

Other than the predicted variable, both familiarisation and test models had the same structure. Both models were fit with Time in trial using orthogonal septic polynomials of the time term up to the fourth order (ot1, ot2, ot3 and ot4), Age (in months) and SES as predictors. Time was nested withing participants interacting with word in the random effects structure. The aim for collapsing across word types and difficulties was to measure overall comprehension effects. A follow up model used looking proportions during test split by word type as the predicted variable. The model was fit with Time in trial using orthogonal septic polynomials of the time term up to the fourth order (ot1, ot2, ot3 and ot4), Age, SES and Word Type (including nouns, verbs and adjectives) as predictors. Time was nested within participants in the random effects structure but in this case we did not include word nor word type since it did not improve the model based on AIC and Anova tests comparing both models.

Results from the model using looking proportions during the familiarisation phase, showed only a main effect of the quadratic Time term (see table 4.5). As can be seen in Figure 4.6, children show overall chance levels of looking indicating no systematic preference for a particular image or side. The negative main effect of the quadratic Time term might be related to negative fluctuations at older ages, however it does not seem consistent enough to suggest consistent biases at test.

Results from the model using looking proportions in overall Comprehension trials during the test phase showed only a main effect of the quadratic Time term (see table 4.6). This indicates that, overall, children increase their looks to the target by the end of the trial. As we can see in Figure 4.7, some children look to target by the end of the trial, particularly older children and higher SES children. These relationships however, are not significant in our model, because we only find a positive main effect of the quadratic Time term. Even though our results show only evidence of an overall effect, some of the possible interactions might be hindered due to the fact that we are analysing overall Comprehension that is, looks to the target in nouns, verbs and adjectives collapsed across three difficulties. This could also be the reason why we see chance performance for many children, because children might find some words hard, which results in less looking towards the target image. Thus, our next step was to account for word type in the Comprehension model, collapsing across difficulties. We did not model Comprehension based on word difficulty because difficulties in India were hard to estimate based only on parental report.

Results from the model using looking proportions in Comprehension trials split by word type (nouns, verbs and adjectives) during the Test phase showed

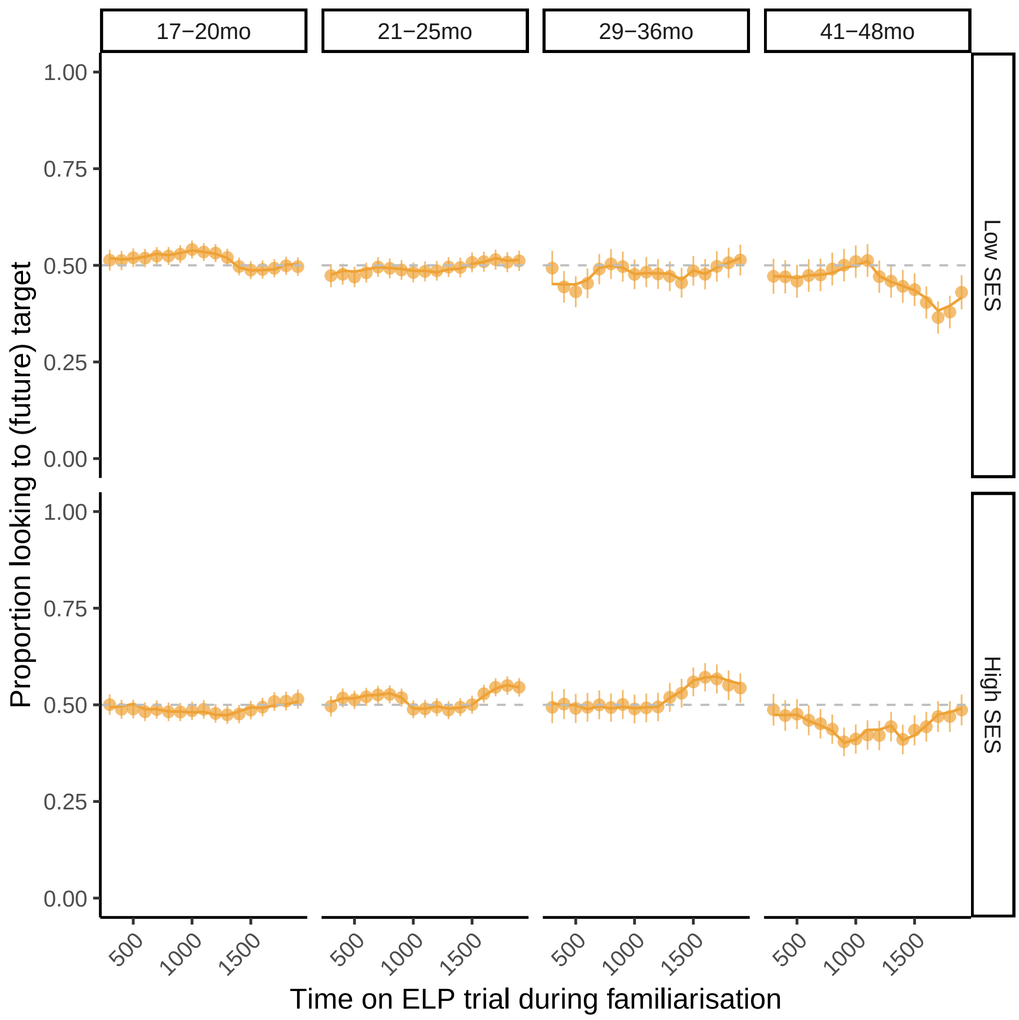


**Figure 4.6:** Model predicted proportion looking to target in Comprehension trials overall by time, Age and SES during familiarisation. Grey dashed line depicts chance performance (0.50). Age in months is split in age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

a main effect of Word Type, several 2 and 3-way interactions including a 2-way interaction between the quadratic Time term and Word Type, 3-way interactions between the linear and quadratic Time terms, Word Type and Age/SES, and 4-way interactions between the linear, quadratic and cubic Time terms, Age, Word Type and SES. Detailed results are presented in Table 4.7). As can be seen in Figure 4.8, younger children do not show a preference towards the target nor the distractor. We see a lot of variability in older children. Older

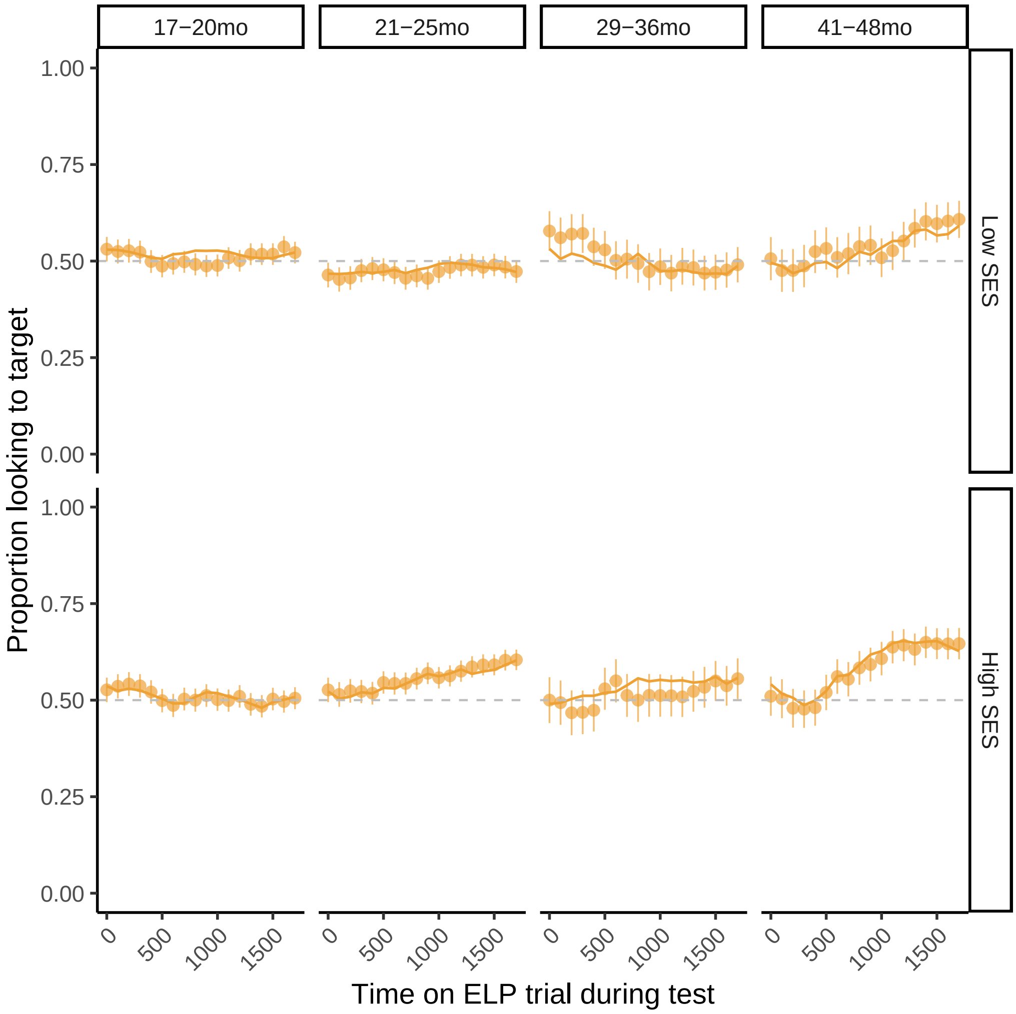


**Figure 4.7:** Model predicted proportion looking to target in Comprehension trials overall by Time, Age and SES during Test. Grey dashed line depicts chance performance (0.50). Age in months was split in age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

children from High SES families look more and quicker to the target for nouns and adjectives, whereas they seem to find verbs harder because they only look (slightly) to the target by the end of the trial. Older children from low SES families seem to struggle with both nouns and adjectives as they show mostly chance-level looking; however, they seem to look more to the target on trials that contain verbs (although it is challenging to interpret this as they also tend to start the test trial looking at the target image on verb trials).

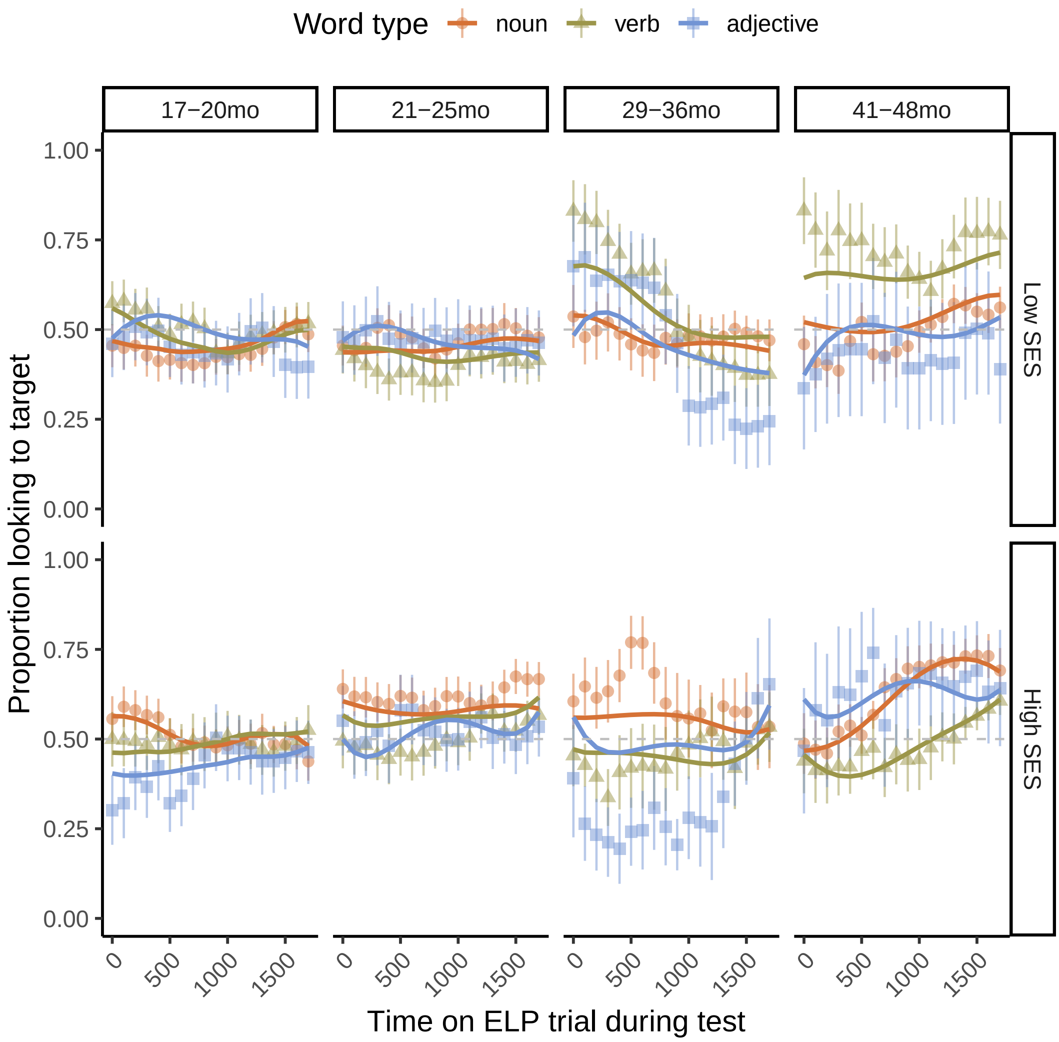


**Figure 4.8:** Model predicted proportion looking to target in Comprehension trials split by Word Type by Time, Age and SES during Test. Grey dashed line depicts chance performance (0.50). Age in months was split in age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

Our results from ELP Comprehension trials at the group level show that older children, particularly those from high SES families, show results that are consistent with findings using the ELP task in the UK with 15 to 27 months old children (see Chapter 2). Once again, however, we see significant looking to the target only at older ages suggesting a delay in early word processing in the India cohort.

Nevertheless, it is possible that individual children are showing better performance earlier on some ELP trials. Here we explored if individual's performance on ELP Comprehension (i.e., proportion looking to the target) is related the their performance on ELP Reinforcement across the three measures extracted from Reinforcement trials: overall accuracy, RT, and the linear time term from the GCA model. We did not find evidence for relationships between ELP Comprehension and any of the ELP Reinforcement measures, although there are positive relationships between Comprehension and Nouns, Verbs and Adjectives since they are part of the same composed measure (see Figure A.12 in Appendix).

**ELP Novelty Bias**

Looking proportions to the novel image for ELP Referent Selection trials during familiarisation (i.e., before the target word was named) were used to measure children's novelty biases and modelled following the GCA approach. On ELP Referent Selection trials, novel images were either paired with familiar images that had previously appeared in Reinforcement trials (highly familiar), or familiar images that had previously appeared in Comprehension trials (familiar). We controlled for the level of familiarity of the familiar image in our model. The model was fit with Time in trial using orthogonal polynomials of the Time term up to the third order (ot1, ot2 and ot3), Age (in months), Familiarity of the familiar image (images from Reinforcement versus from Comprehension), and SES as fixed effects. Time was nested within participants interacting with image familiarity in the random effects structure.

Results from the model revealed a 3-way interaction between the cubic Time term, Familiarity of the familiar image and SES, as well as a 4-way interaction between the cubic Time term, Age, Familiarity of the familiar image and SES (see table 4.8 placed at the end of the chapter). As it can be seen in Figure 4.9, the younger children from our sample show overall chance levels indicating no systematic preference for a particular image. Low SES young children, however, show a familiarity bias for highly familiar images (those that appeared also during Reinforcement trials) by 21-25 months of age; high SES children show a similar bias but it is weaker. This pattern is reversed at older ages. Both high and low SES older children in our sample looked more towards the novel image only when this appeared in the context of a highly familiar image (i.e., an image from Reinforcement trials). Some of the older high SES children also look more towards the novel by the end of the trial when paired with familiar images (i.e., those that appeared in Comprehension trials).

Next, we examined novelty biases at the individual level. We did not find any relationships between novelty biases and performance on ELP Reinforcement or ELP Comprehension, neither between Familiarity Biases (in which the familiar image also appeared in comprehension trials) and ELP Reinforcement nor ELP Comprehension. However, we see a positive significant relationship between looks to the novel and looks to the familiar pre-test (Novelty Bias and Familiarity Bias Reinforcement $t = 4.558$; $r = 0.489$; $p < .001$ and Novelty Bias and Familiarity Bias Comprehension $t = 7.973$; $r = 0.700$; $p < .001$) which might indicate that those children are more on task (see full set of r values in Figure A.13 in Appendix).

### ELP Referent Selection

Looking proportions to the target for ELP Referent Selection trials during test were modelled following the GCA approach. The model was fit with Time in trial using orthogonal polynomials of the time term up to the third order (ot1,

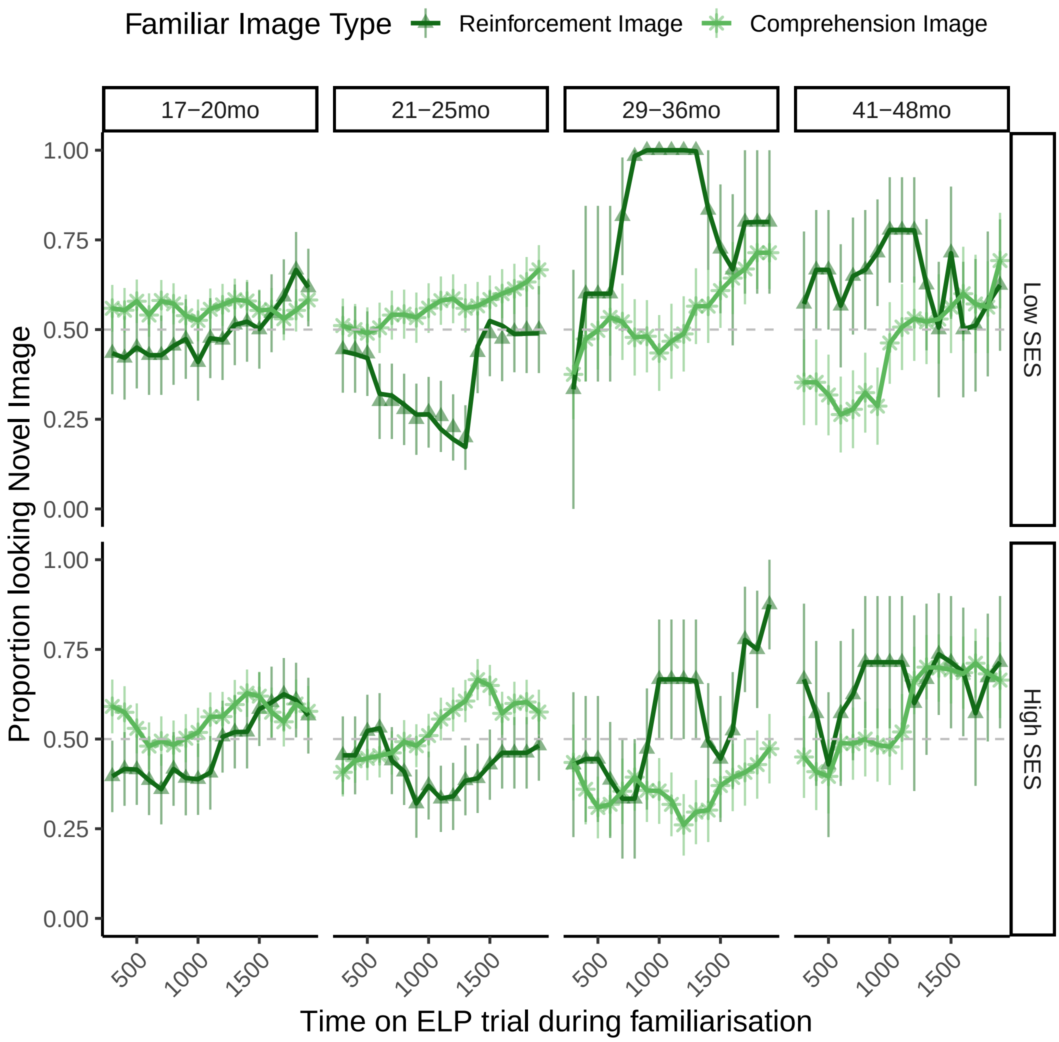


**Figure 4.9:** Model predicted proportion looking to the novel image in Referent Selection trials during familiarisation, split by Familiarity of the familiar image, Time, Age and SES. Grey dashed line depicts chance performance (0.50). Age in months was split in age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

ot2 and ot3), Age (in months), Word Type which indicates if the target was the novel or the familiar/noun, and SES as fixed effects. Time was nested within participants in the random effects structure.

Results from the model revealed a main effect of Word Type (novel versus familiar), some 2 and 3-way interactions, and a 4-way interaction between the cubic Time term, Age, Word Type and SES (see Table 4.9 placed at the end of the chapter for full model results). As can be seen in Figure 4.10, younger children are mostly at chance levels, showing no preference for the target nor the distractor. This is the case for trials in which the novel was the target and trials where a familiar noun was the target. The youngest children in our sample from low SES seem to have a preference for the familiar image, because they look more to the target in trials were the target was a familiar noun, and look more to the distractor in trials were the target was a novel word-object. Older children look more to the target overall. Older children from high SES seem to be particularly good at recognising the target when this was a familiar noun. Overall, we see quite different patterns across word type at different ages as well as across the SES groups as reflected in the significant 2 and 3-way interactions including the linear, quadratic, cubic and quartic Time terms. We also see different looking trajectories across children as a function of age, SES and Time. However, a recurrent pattern is that most children check both target and distractor during the trial. This behaviour might be reflected in the 4-way interaction that we find only with the cubic Time term.

At the individual level, we explored possible relationships between Referent Selection trials and other ELP measures using correlation analysis. We split Referent Selection trials by word type and calculated the mean proportion of looks to target from word onset to 1800 ms for trials where the target was novel and for trials where the target was familiar. We found a positive correlation between ELP Referent Selection trials in which the target was familiar and ELP Comprehension trials but significance goes away when correcting for multiple comparisons ($t = 2.11$; $r = 0.251$, $p = 0.038$). This relationship indicates that children who are good at recognising familiar words during ELP Comprehension are also good at recognising the familiar target on Referent

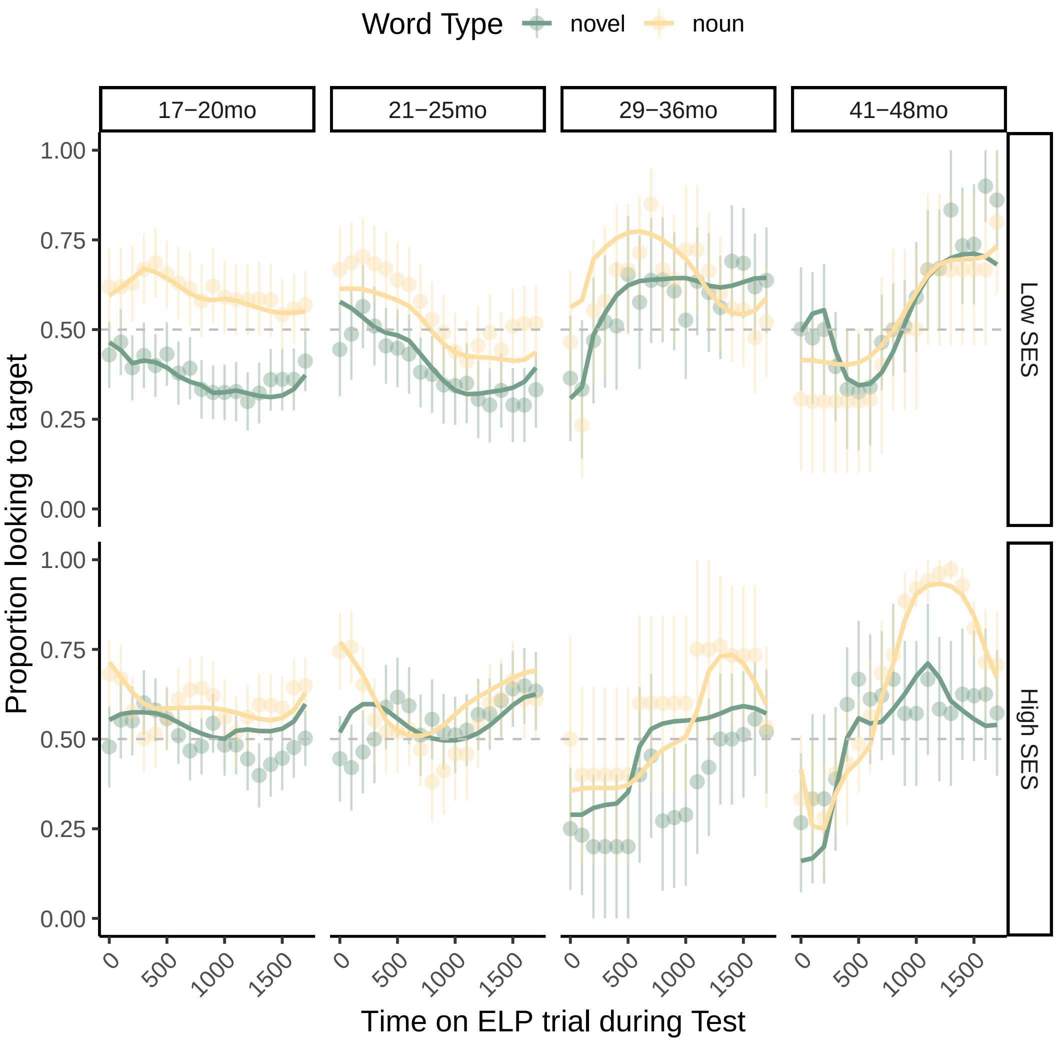


**Figure 4.10:** Model predicted proportion looking to the target in the test phase of Referent Selection trials. The model includes proportion looks to target by Time, Age in months, Word Type (familiar versus novel) and SES. Grey dashed line depicts chance performance (0.50). Age in months was split in age groups to facilitate visualization. Different colours show target word types, in yellow looks to the target when that was a familiar noun, in green looks to the target when that novel. Dots indicate the raw mean data per each 100 ms time bin including standard deviation. Lines show the model predictions.

Selection trials.

### ELP Retention

Looking to the target on ELP Retention trials both during familiarisation and test were modelled following the GCA approach. Note that we used one model where the predicted variable was looking proportions to the target during familiarisation (i.e., the image that will become the target during test), and another model were the predicted variable was looking proportions to the target during test. Both models were fit with septic polynomials of the time term up to the third order (ot1, ot2 and ot3), Age (in months), and SES as fixed effects. Time was nested within participants in the random effects structure. In these models, we only included trials were children looked more than 50% at the target on Referent Selection trials. This was to make sure that children were paying attention to the correct image when they learned the label for the novel object. This meant that a total of 103 out of 177 children were included in the model using looking proportions during the familiarisation phase of Retention trials, and a total of 81 out of 177 children were included in the model using looking proportions during the test phase of Retention trials.

Results from the model using looking proportions during the familiarisation phase showed a 2-way interaction between the linear Time term and SES, and a 3-way interaction between the linear Time term, Age and SES (see Table 4.10. As can be seen in Figure 4.11, children from high SES families seem to oscillate between both images. Older children from low SES families seem to look more to the image that will be the target at the end of the familiarisation trial. It is hard to know if this pattern is related to naming effects during Referent Selection trials. That is, children were previously exposed to both novel images during Referent Selection and one of them was named. It is possible that children remember which image was named and thus, show a preference during familiarisation. However, if that were the case, we might

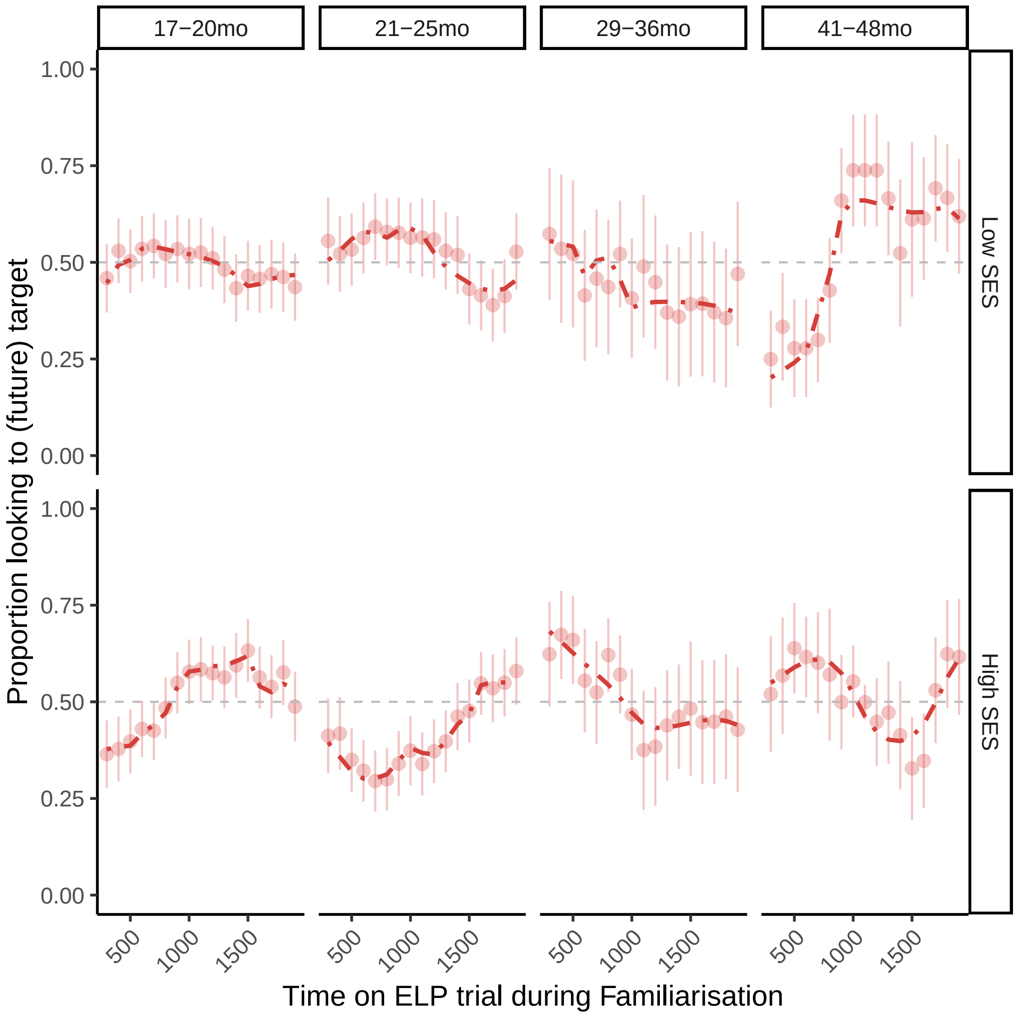


**Figure 4.11:** Model predicted proportion looking to target in Retention trials Time, Age and SES during familiarisation. Grey dashed line depicts chance performance (0.50). Age in months is split in age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

expect the preference to look at the "future" target to be consistent across children from low and high SES backgrounds; this is not what we see in our data.

Results from the model using looking proportions during the test phase showed significant main effects of Age and SES, as well as a significant 2-way interaction between Age and SES (see Table 4.11). As can be seen in Figure 4.12, children are generally at chance levels, meaning they do not show

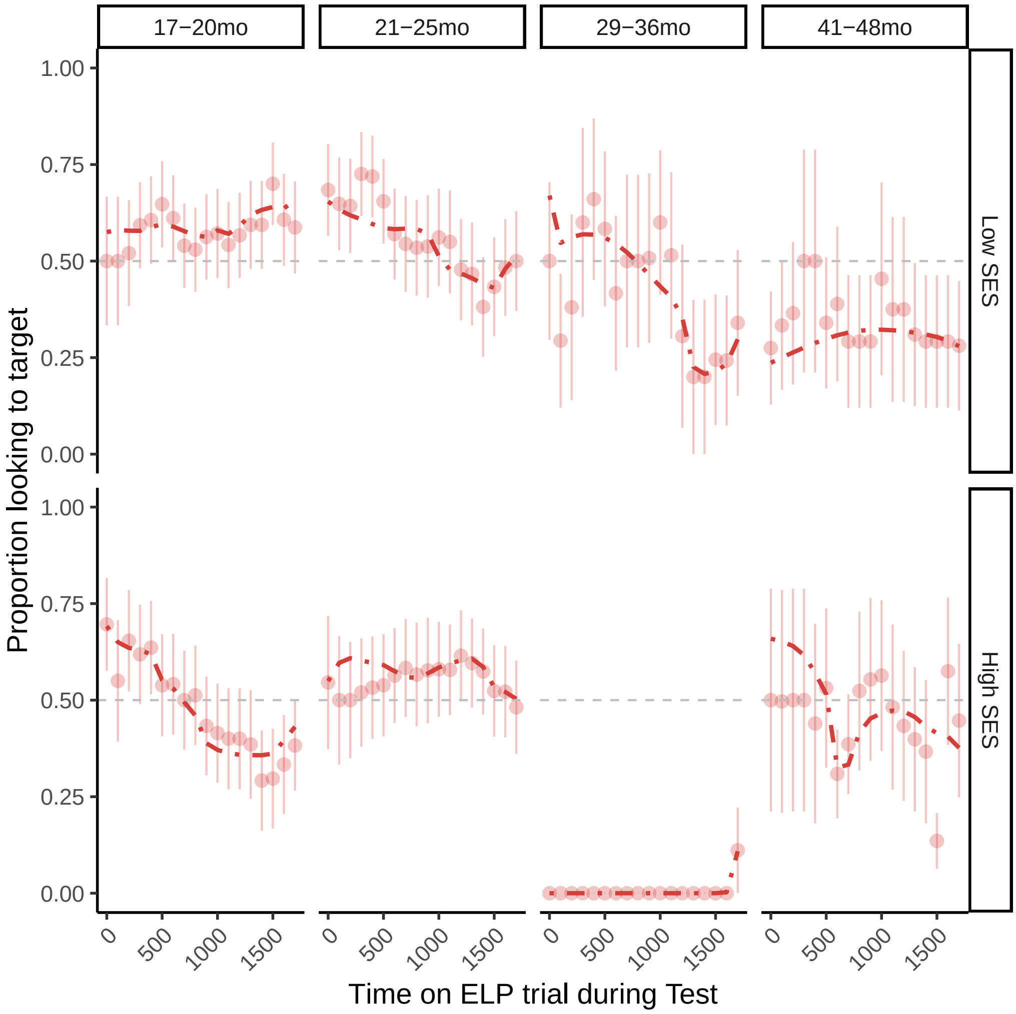


**Figure 4.12:** Model predicted proportion looking to target in Retention trials Time, Age and SES during test. Grey dashed line depicts chance performance (0.50). Age in months is split in age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

retention of the new word-object mappings. Indeed, older, low SES children tend to look to the distractor. Critically, we see high variability in our data. This might be due to the smaller sample included in this analysis. This can be seen in missing data for some ages which lowers the power of our analysis.

At the individual level, we explored possible relationships between Retention trials and other ELP measures using correlation analyses. We found significant positive correlations between individual's performance on ELP Re-

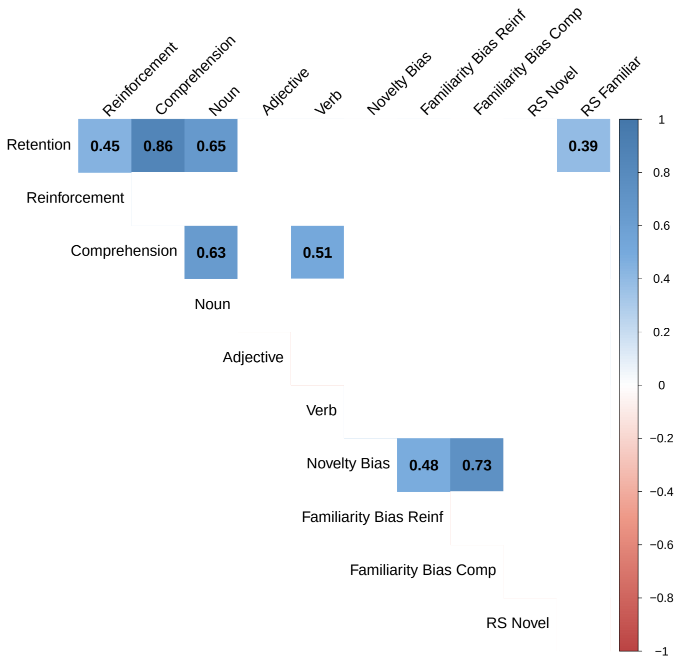


**Figure 4.13:** Correlation matrix for ELP Retention, ELP Reinforcement, ELP Comprehension and ELP Referent Selection split by trials were the target was novel (RS Novel) and familiar (RS Familiar). Only correlations smaller than 0.01 (sig.level <0.01) are displayed. Correlations were performed only on those children who had data for all the subtasks in this section. Positive correlations are in blue and negative correlations in red, the strength of the colour indicates the strength of the relationship.

tention trials and accuracy on ELP Reinforcement trials (t = 3.867, r = 0.446; p <.001). This indicates that children with high proportions of looks to the target on Retention trials also have high proportions of looks to target on Reinforcement trials. We also found a significant positive relationship between performance of ELP Retention trials and ELP overall Comprehension (t = 12.805; r = 0.855; p <.001), this relationship holds for nouns when splitting comprehension by word type (t = 6.627; r = 0.650; p <.001). There is also a significant correlations between Retention and Referent Selection on familiar trials (t = 3.291; r = 0.391, p <.001). These results can be seen in Figure 4.13

(see full set of r values on Figure A.14 in Appendix).

These results are particularly interesting because we do not find evidence of word retention at the group level. However, individual children's performance on Retention trials is correlated across several ELP measures.

### 4.3.2 ELP 30-months-old Retest

Looking proportions to the target across ELP trials for the retest data were modelled using the GCA approach. The model was fit with Time in trial using orthogonal septic polynomials of the time term up to the fourth order (ot1, ot2, ot3 and ot4), Trial Type, Age (in months) and SES as fixed effects. Time and Participants were also added as random effects. Since this was a retest of children who had done ELP at younger ages, we wanted to measure if the task was still sensitive to differences in performance across the language processing measures (or ELP trials) in children who already had experience with it, before looking at relationships across test and retest. Note that this is the only model in this section that includes looking data from the 30-month-old group alone; all other models measure relationships across looking data in the same children at both time points (test, retest).

Results showed main effects of the linear and quadratic Time terms, Age, Trial Type and SES, as well as multiple interactions among the variables including a 3-way interaction between the linear and quadratic Time terms and Age, Trial Type and SES. A full report of the results can be seen in Table 4.12, at the end of this chapter.

The model results are plotted in Figure 4.14. As can be seen, the 30-months-old retest group show age effects, primarily for high SES children. Those children showed more looks to target for Reinforcement and Referent Selection trials, and more chance levels of looking for Comprehension and Retention trials. Interestingly, younger low SES children also show more looks towards the target on Reinforcement trials, however, they are at chance levels at older ages. Greater looks to the target on Reinforcement and Referent Selection trials are also seen in older children in our 18-month-old Group data.

These results indicate strong SES effects in our retest as well as differences across trials (or ELP measures) in our Indian population at retest.

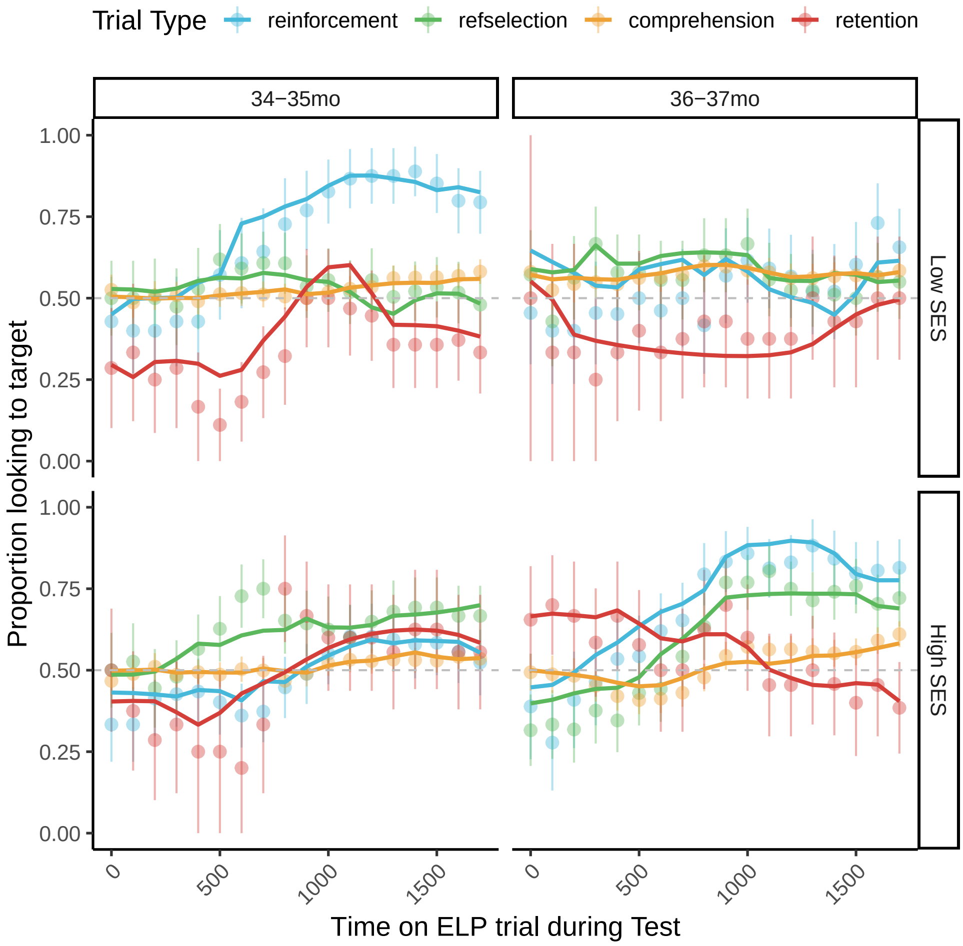


**Figure 4.14:** Model predicted proportion looking to target by Trial Type by Age and SES. Grey dashed line depicts chance performance (0.50). Age in months is split in age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

Since the overall model showed evidence of differences across ELP trial types as well as SES, we modeled each trial type separately in more detail.

We used the previous model and adapted it as necessary. Due to the small sample size, some of these individual models lacked enough statistical power to extract robust conclusions. We report these in Appendix A.

### 4.3.3 Relationships between Test-Retest

In this section, we address Q3, that is, how is ELP performance at test related to performance at retest at the group level? We then address Q4, how is ELP performance at test related to performance at retest at the individual level? Note that here we only report looking data during the test phase of the trial. We did not look at the familiarisation phase because we did not find consistent evidence of biases in the test data. Mimicking the procedure in Chapter2 we only explored relationships between ELP Reinforcement, ELP Comprehension and ELP Reinforcement between Test and Retest.

Looking proportions to the target only for ELP Reinforcement trials at both test and retest were modelled following the GCA approach in a single model. The model was fit with septic polynomials of the Time term up to the second order (ot1 and ot2), Age (in months), Repetition Pair that accounted for the number of repetitions of the Reinforcement image pair, SES and Test Type (test, retest) as fixed effects. Both SES and Test Type were scaled and centered. Time was nested within participants in the random effects structure interacting with Test Type.

Results from the model, showed a main effect of the linear Time term, Age, Repetition Pair Count and SES. There were several 2, 3 and 4-way interaction as well as a 5-way interaction between the linear and quadratic Time term and Age, Repetition Pair Count, Test Type and SES (see Table 4.13 for full model details).

As can be seen in Figure 4.15, all children look significantly more to the tar-

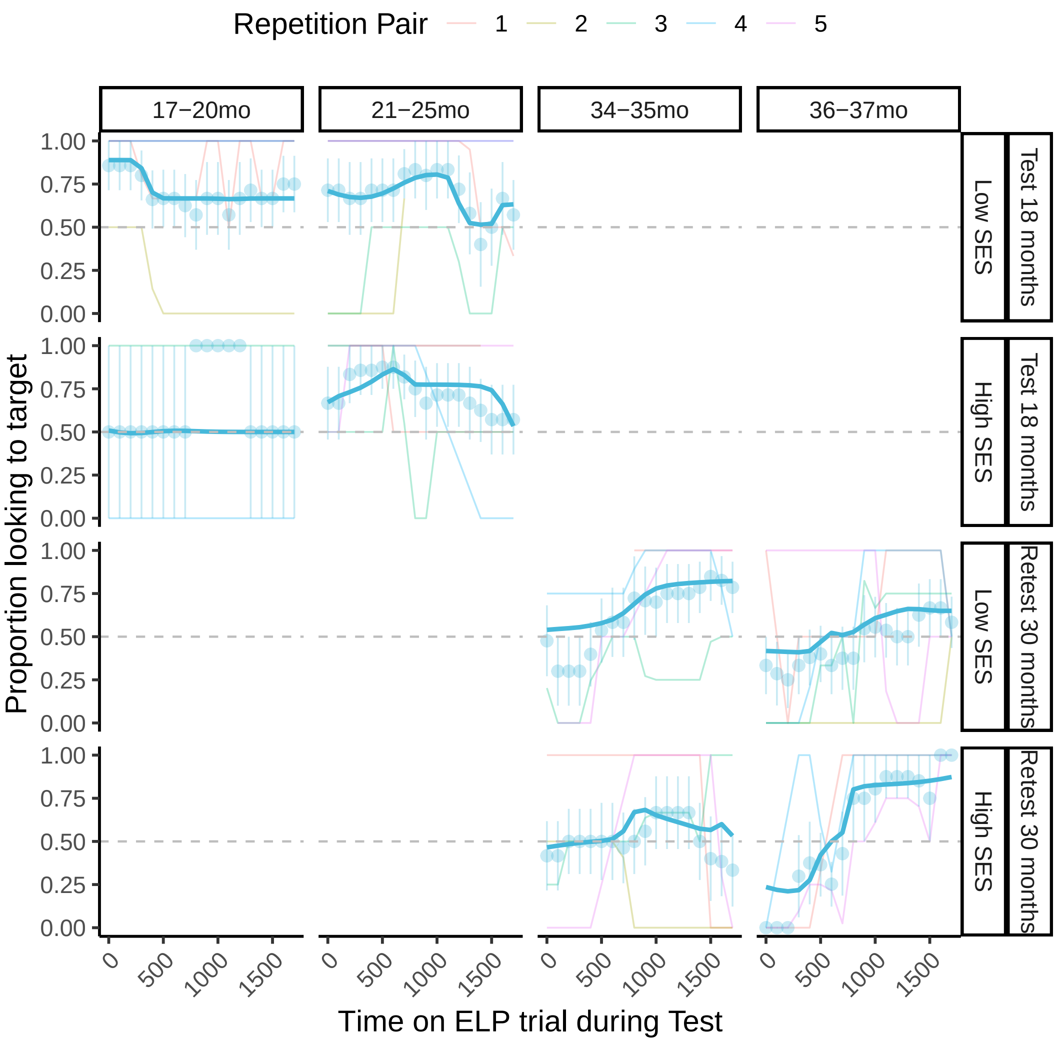


**Figure 4.15:** Model predicted proportion looking to target in Reinforcement trials by Age, Repetition Pair Count and SES, split by Test Type, test at 18 months versus retest at 30 months. Grey dashed line depicts chance performance (0.50). Age in months is split in age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

get; however, due to the small sample size, we did not have much data in some groups (e.g., younger children from high SES families at the 18-months-old test). Thus, results should be interpreted with caution. We generally found an increase in looking to the target at retest, particularly with the high SES group. Repetition pair seems to also play a role with children looking more to the target when exposed to more more repetitions. The 5-way interaction reveals that children look more to target as a function of Age, SES, Repetition Pair Count and SES. This is reflected in older children, particularly those from high SES families, looking more and more quickly to the target.

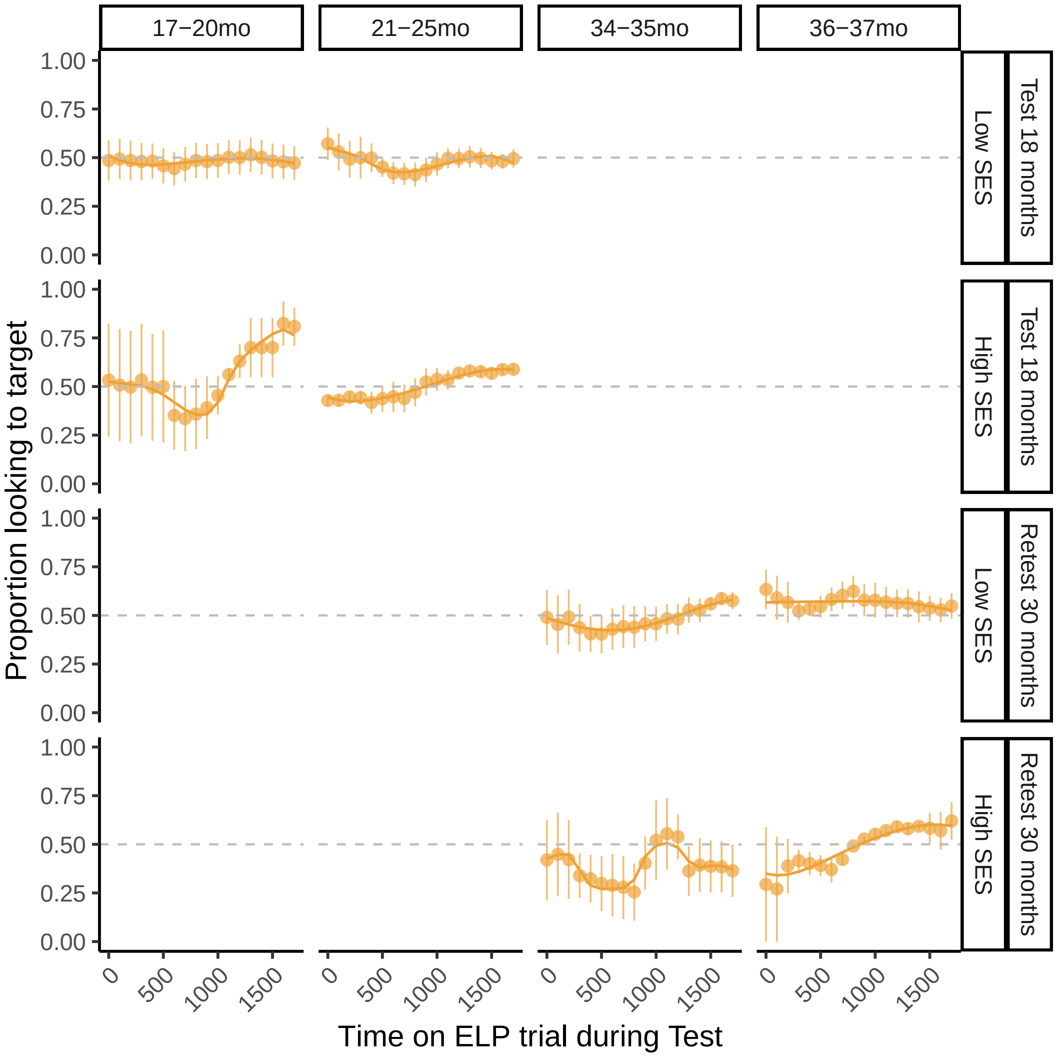


**Figure 4.16:** Model predicted proportion looking to target in overall Comprehension trials by Age, SES, and split by Test Type, test at 18 months versus retest at 30 months. Grey dashed line depicts chance performance (0.50). Age in months is split in age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

Looking proportions to the target for ELP Comprehension trials at both 18-months test and 30-months retest were modelled following the GCA approach. The model was fit with septic polynomials of the Time term up to the third order (ot1, ot2 and ot3), Age (in months), SES and Test Type (test, retest) as fixed effects. Both SES and Test Type were scaled and centered. Time was nested within participants in the random effects structure interacting with Test Type.

Results from the model show no significant effects (see Table 4.14 for full model details). As can be seen in Figure 4.16, children were at chance levels for both test at 18-months and retest at 30-months indicating no preference for target nor distractor when trials were collapsed across nouns, verbs and adjectives from different difficulties. A follow up model splitting by word type was explored but, due to the small sample size, a lot of data was missing across ages and SES; thus, it was hard to extract conclusions from the results of that model (details of that analysis are attached in Appendix A, Figure 4.16 plots the data with model predictions, results from the model can be seen on Table A.4).

We also modeled looking proportions to the target for ELP Retention trials at both 18-months test and 30-months retest. For consistency with previous analysis, we only included trials were children looked more than 50% at the target on Referent Selection trials. Only 12 children were included in this model This makes it hard to extract clear conclusions from this analysis (details of that analysis are attached in Appendix A, Figure A.16 plots the data with model predictions, results from the model can be seen on Table A.5).

Finally, we explored relationships between individuals' performance using correlational analysis. We did not find any significant relationships, likely due to our low power.

## 4.4 Discussion

In this study we measured language processing abilities in children from rural India learning Awadhi. In order to do this, the first aim of this study was to adapt the ELP task that we developed in Chapter 2, to a different language, population and culture. The second aim was to gather measures of language processing in Indian children at different ages while taking SES into account. Finally, the third aim of this project was to capture the developmental trajectory of early word learning; thus, we administered the ELP to the same group of children at two different time points (18 and 30 months of age). Here we also controlled for SES in our sample.

Overall, this project shows that is possible to design and adapt culturally valid measures in rural non-western settings. We did face a lot of obstacles in this process, from lost of power to daily linguistic and cultural barriers. A crucial element in this study was the close collaboration with our partners from the CEL in India. Their team is very diverse and included local people from Shivgarh. They were truly essential to adapt ELP to the site and administer it to the children. Working together with them is what made possible to collect such a large number of participants from diverse SES backgrounds. Our results highlight that ELP generalises to the Indian population. However, only older children show good performance which could indicate a delay in language ability. We find strong SES effects, with lower SES children performing worst than higher SES children at most ELP measures. At the individual level we find relationships between retention abilities and word recognition as well as disambiguation skills.

### 4.4.1 Possible delay in language processes

Our results show that we were able to successfully adapt a new measure to a different language, population and culture. At the group level, the ELP task is able to capture different language processes involved in word learning. Young children in India seem to struggle with the ELP task, but older children in India show differences in looking patterns across the five ELP measures. This indicated that, at those older ages, the task is sensitive enough to capture different processes involved in word learning. Moreover, in most of the ELP measures older children in India replicate similar looking patterns to those shown by western children (A. Fernald et al., 1998, 2001; Bion et al., 2013). We did not find clear evidence of biases during the familiarisation phase of ELP, with the exception of the novelty bias measures which was designed to capture novelty biases. This then indicates that the processing differences seen in the task reflect children's performance on the task based on their abilities, rather than stimulus-driven effects.

Our results show that SES effects can already be see as early as 17 months and are very pronounced during toddlerhood. This study also shows that SES based on maternal education is a relevant construct in rural India. In this context, maternal education reflects access to school and basic literacy and thus, it is an index of poverty. This highlights the role of the primary caregiver, which is usually the mother in early infancy in rural India, in children's early language development and later abilities, as proposed by some studies (Hoff, 2003). As well as the effect of poverty in children's language skills during childhood and toddlerhood.

We also find a lot of variability in the data, some is due to the SES differences. Children with mothers that are more highly educated perform better at ELP (i.e. they are faster and more accurate). These differences can already be seen as early as 17 months in some ELP measures such as Referent Selection, but they are clearly salient after 25 months. This is partly because younger children are mostly performing at levels not different from chance, making it is hard to know if SES differences across the other ELP measures are also present early on. What is specially interesting is that this fits the same age ranges reported in previous literature with western samples. A. Fernald et al. (2013) found that already at 18 months, children showed significant disparities in vocabulary and language processing efficiency based on SES. Those differences across higher and lower SES families become critical for language development by 24 months, with low SES children showing a 6-month gap in speed of word processing compared to their higher SES peers (A. Fernald et al., 2013). It is hard to know if our Indian children show a 6 months gap with our current data, mainly because we always used age as a continuous variable in our models, and also due to the variability in our data. However, it is possible that the biggest SES differences in our sample are at older ages, which is when we see the biggest differences in word recognition abilities. This could be because ELP in India is more sensitive at older ages when performance differs from chance levels. However, it could also reflect the cumulative effect of SES on language, since SES has been reported to have long term effects up until adulthood (e.g., Pakulak & Neville, 2010).

Both age and SES effects can be seen at the group level in all the ELP measures. On ELP Reinforcement, older children's performance, particularly those from High SES families, is similar to findings from previous literature measuring speed of word processing in western samples (A. Fernald et al., 1998, 2001). Note, however, that we do not see robust looking to the target with high SES Indian children until 29-36 months; thus, it appears Indian children's speed of processing does not reach the levels seen in Western samples until later in development, even in the high SES cohort. However, caution is warranted as these results could also reflect a lack of familiarity with the testing conditions. Most children in our sample from India had very little experience with TV screens, something that is very uncommon in households from that rural area. That said, the children in our Indian sample were also part of a longitudinal study that tracked their development from 6 to 36 months of age. Once a year, these children spent a day or two in the Laboratory set up in Shivgarh doing several tasks measuring attention, visual working memory, dyadic interaction and motor skills. Some of those tasks also involved screens and eye-tracking techniques. This means that when children did ELP at 18 months, they all had several experiences with screens as well as with the experimental set up. While this is not equivalent to the extensive exposure to screens that children from Western households are likely to have experienced, it does somewhat mitigate against the concern that their performance is solely due to a lack of familiarity with the experimental setup.

Similar age and SES effects have been seen on ELP Comprehension across word types at the group level. Even though younger children are at chance levels, older children from high SES households are able to quickly identify the referent of nouns and adjectives. They did struggle, however, with verbs. This is a very similar pattern to what has been previously reported in the UK using the same task (see Chapter 2). It is possible that we see lower performance with verbs because actions might be harder to recognise using 2-dimensional images, in comparison to nouns and adjectives. We also see a lot of variability in older children from low SES households who seem to struggle with both nouns and adjectives as they show mostly chance-level looking. This is interesting because they seem to be able to recognise quite well highly familiar nouns that appear on ELP Reinforcement. This could be due to two reasons: first, it suggests that in India, our selection of highly familiar nouns is appropriate. When we designed ELP, difficulty was established based on word frequency on children's natural input. We used a set of corpora from British infants to measure that. Those frequencies were matched with Indian mum's reports about frequency of word occurrence. This indicates that Indian mums were very accurate at reporting highly familiar words in their children's input. Second, it is possible that children are better at our Reinforcement words because they repeat several times in the task. Thus, alternatively, children learned those words during ELP.

Contrary to the two previous measures where we see the stronger SES effects at older ages (e.g., ELP Reinforcement), ELP Novelty bias shows that SES effects are present at earlier ages in our sample. By 21-25 months of age, low SES children in our sample show a familiarity bias, but only for highly familiar images (those that appeared during Reinforcement trials). This is interesting because low SES 21-to-25 month old children do not show evidence of recognising the target on reinforcement trials (i.e., they are at chance levels). However, when presented with those familiar images in the context of a novel one (before being prompted with the target word), low SES 21-to-25 month old children prefer to look at the familiar image that they have already seen several times during Reinforcement trials compared to the novel image. Thus, showing a familiarity bias. This is not the case for familiar images that were presented only once such as those that appeared on Comprehension trials. This suggests these children are remembering which images they have seen before in the task. Similar learning is also seen when low SES 21-to-25 months show looks to the target on the last repetition of Reinforcement trials. This familiarity bias is less pronounced for the same age high SES children, who show a familiarity bias on trials where the novel image was paired with a familiar image that appeared during Reinforcement trials, but a novelty bias on trials where the novel image was paired with a familiar image that appeared during Comprehension trials. This might suggest that higher SES 21-25 months are able to suppress the familiarity bias and look at the novel image.

An interesting finding related to novelty biases, is that children's responses change with age. The pattern we just described among high and low SES 21-25-months old is reversed at older ages. Both high and low SES older children in our sample looked more towards the novel image only when it appeared in the context of a highly familiar image (i.e., an image from Reinforcement trials). This indicates that, at older ages, highly familiar images in the context of a novel ones push attention towards novelty. Some of the older high SES children also look more towards the novel image by the end of the trial when paired with familiar images (i.e., those that appeared in Comprehension trials). This suggests that high SES older children have a better memory of the images from the Comprehension trials leading to stronger attraction to novelty. Age effects in relation to novelty biases have been reported in western children (Kucker et al., 2018; Horst & Samuelson, 2008). Although in those studies older children are less attracted to novelty than younger children, in our task more attraction to novelty might be a good thing. On ELP, children are constantly prompted to look at familiar images during all the task, it is possible that this reduces children's attraction to novelty in our trials. Less attraction to novelty though, could mean a worse performance on reference selection and retention trials. Thus, in this task, novelty biases might be a good thing because they would allow children to suppress familiarity biases and be better at disambiguation on Referent Selection trials.

Results on ELP on Referent Selection trials also showed an age effect. Older children in our sample showed better disambiguation skills. Older children from high SES were particularly good at recognising the target when this was a familiar noun. When considered together with the novelty bias results, this suggests that children in India show weaker attention to novelty that western children (Kucker et al., 2018; Horst & Samuelson, 2008). It is possible that novelty biases are weaker because all the objects are relatively novel as children in India rarely look at objects on a video monitor. This could be tested in the Indian sample by doing the RSR task with real objects as in Horst and Samuelson (2008). This would allow to assess if familiarity and novelty biases in this population are related to the ELP task or the setup.

At the group level, the children in our Indian sample did not show evidence of retention of the new word-object mappings that may have been formed during the referent selection trials. In western samples, children are reported to be able to remember new word-object associations learned during referent selection by 30 months of age (Bion et al., 2013). This has been reported using a looking task with 2-dimensional images similarly to ELP. However, even the 41- to 48-month-old children in our Indian sample do not show evidence of remembering those associations. In fact, if we consider our results for both ELP Referent Selection and Retention, the Indian children in our sample who were between 29-36 months look similar to 18-month-old Western children in Bion et al. (2013). Likewise, the 41- to 48-month-old children in our Indian sample show similar performance to Bion et al. (2013) 24 month-old children (particularly the higher SES group. This might indicate a possible delay in their development and thus, explain what we do not see evidence of retention in the older children of our sample. It is also the case that we have fewer Retention trials in our sample, suggesting that children found them hard and did not contribute data, which might have lowered our statistical power.

Our last study question aimed to look at individual trajectories between test (at 18 months) and retest (at approximately 30 months). At the group level we see very similar patterns that those found at test. Unfortunately, due to the low sample size, there is some noise in our analyses and we and not able to capture any relationships across test and retest. Another issue is that in out first analysis we see that children at 18 and 30 months, the approximate ages of our test and retest, many children are at chance. So it is harder to capture individual differences across two groups that are mostly at chance. This shows how hard can be to test children in rural India.

### 4.4.2 Individual children can remember newly learned words

A particular feature of ELP is that integrates several language processing measures in a single task. Thus, one of our questions was how individual children performed across ELP measures. It is interesting that even though many children are at chance levels, particularly the younger children, when considering all participants across ages at the individual level, we find relationships across some ELP measures, particularly for word retention. A highlight of our results is that even though as a group children performed at chance levels in Retention trials, we found relationships between Retention and online word comprehension measures (Reinforcement and Comprehension) as well as between Retention and referent selection measures (for both novel and familiar targets) at the individual level. This indicates that individual children who had good comprehension abilities or were better at recognising highly familiar words, were also good at remembering new word-object associations. Similarly, children who were good at disambiguation, both when they had to find the novel target as well as when they had to find the familiar one, had a good performance on Retention. This is expected because to be able to find the target or retention trials, children had to pay attention to the correct target on referent selection trials. These findings are in line with previous studies with Western children suggesting that good lexical skills help children learn new words and that good disambiguation abilities help children retain new words (Bion et al., 2013; Kucker et al., 2020; Lany, 2018). Importantly, we see this relationship in children from both low and high SES families and when collapsing across ages. This indicates that this effect is robust and can be seen in individual children from different SES backgrounds.

### 4.4.3 Limitations and Future directions

Our study might benefit from extending some of our analyses. For example, we aimed to measure disambiguation abilities and thus, we only looked the trials were the target was familiar versus novel. Since we see the role that repetition plays in novelty and familiarity biases at 21-25 months in Indian children, it would be interesting to look at Referent Selection performance while taking into account whether the familiar image appeared in Reinforcement versus in Comprehension trials. Some literature shows that repetition facilitates word learning (Twomey, Ranson, & Horst, 2014; Axelsson & Horst, 2014), thus it is possible that children show better disambiguation skills in trials were the novel image was paired with a highly familiar image (i.e., those that appeared also in Reinforcement trials), which repeat several times during ELP. This is particularly interesting given the effects of familiarity we see on the novelty bias measures for which we did this distinction.

An additional set of model could also explore relationships between more than two ELP measures. Accounting for several ELP measures could help understand more in detail the relationships among different language processes (e.g., novelty biases, referent selection and retention abilities). Moreover, this would allow to assess the directionality of the effects between ELP measures.

Finally, future work should also take into account other cognitive measures in this sample, so we could investigate the impact of poverty in different aspects of development which are likely to influence each other. Moreover, additional variables should be taken into account when measuring language development such as the child previous experience with language (e.g. the amount of input they are exposed to). Future work looking at those dimension would help better understand how early language experiences influence later language skills and cognitive abilities.

## 4.5 Conclusion

In conclusion, the present study provides evidence that it is possible to efficiently adapt a portable language task cross-culturally. Moreover, integrating multiple measures provides a unique scope to the language processing abilities in children. In this study, we were able to measure and related speed of processing, online comprehension, novelty biases, referent selection and retention abilities in a large sample of children growing up in a remote location. We were also able to capture the effect of poverty though SES, which shows strong effects in our sample from early on and across development.

The next step is to look at the relationships between early language processes and children's previous home experiences with language. This would allow to investigate the role of language input in a culture where families hugely diverse from what has been previously studied in western samples.

These results set the stage for future work to measure language processing abilities in infancy in order to predict longer-term language and cognitive outcomes, as well as working to understand how changes in children's environment lead to differences in language processing abilities over development. Importantly, understanding the mechanisms that underlie these relationships could provide empirical evidence that inform intervention efforts early in development.

## 4.6 Significance Tables

**Table 4.2:** Regression results for ELP at 18 months across all trials during Test

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 402.00 | 1.00 | 0.00 | *** |
| ot1 | 10.95 | 1.00 | 0.00 | *** |
| ot2 | 44.32 | 1.00 | 0.00 | *** |
| ot3 | 0.17 | 1.00 | 0.68 | |
| ot4 | 22.28 | 1.00 | 0.00 | *** |
| Age | 144.09 | 1.00 | 0.00 | *** |
| TrialType | 1228.34 | 3.00 | 0.00 | *** |
| SES | 147.29 | 1.00 | 0.00 | *** |
| ot1:Age | 3.66 | 1.00 | 0.06 | . |
| ot2:Age | 16.81 | 1.00 | 0.00 | *** |
| ot3:Age | 3.98 | 1.00 | 0.05 | * |
| ot4:Age | 28.45 | 1.00 | 0.00 | *** |
| ot1:TrialType | 168.89 | 3.00 | 0.00 | *** |
| ot2:TrialType | 82.88 | 3.00 | 0.00 | *** |
| ot3:TrialType | 13.71 | 3.00 | 0.00 | ** |
| ot4:TrialType | 84.01 | 3.00 | 0.00 | *** |
| Age:TrialType | 1040.14 | 3.00 | 0.00 | *** |
| ot1:SES | 1.29 | 1.00 | 0.26 | |
| ot2:SES | 0.25 | 1.00 | 0.62 | |
| ot3:SES | 1.88 | 1.00 | 0.17 | |
| ot4:SES | 13.92 | 1.00 | 0.00 | *** |
| Age:SES | 181.91 | 1.00 | 0.00 | *** |
| TrialType:SES | 343.83 | 3.00 | 0.00 | *** |
| ot1:Age:TrialType | 215.21 | 3.00 | 0.00 | *** |
| ot2:Age:TrialType | 67.32 | 3.00 | 0.00 | *** |
| ot3:Age:TrialType | 19.31 | 3.00 | 0.00 | *** |
| ot4:Age:TrialType | 89.53 | 3.00 | 0.00 | *** |
| ot1:Age:SES | 4.29 | 1.00 | 0.04 | * |
| ot2:Age:SES | 0.56 | 1.00 | 0.45 | |
| ot3:Age:SES | 13.04 | 1.00 | 0.00 | *** |
| ot4:Age:SES | 28.30 | 1.00 | 0.00 | *** |
| ot1:TrialType:SES | 53.18 | 3.00 | 0.00 | *** |
| ot2:TrialType:SES | 113.42 | 3.00 | 0.00 | *** |
| ot3:TrialType:SES | 18.15 | 3.00 | 0.00 | *** |
| ot4:TrialType:SES | 72.12 | 3.00 | 0.00 | *** |
| Age:TrialType:SES | 356.82 | 3.00 | 0.00 | *** |
| ot1:Age:TrialType:SES | 122.22 | 3.00 | 0.00 | *** |
| ot2:Age:TrialType:SES | 181.50 | 3.00 | 0.00 | *** |
| ot3:Age:TrialType:SES | 45.69 | 3.00 | 0.00 | *** |
| ot4:Age:TrialType:SES | 80.93 | 3.00 | 0.00 | *** |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic), ot3 (cubic), ot4 (quartic), Age in months, TrialType including all ELP trials and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 4.3:** Regression results for Reinforcement Trials at 18 months during Familiarisation

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 2.08 | 1.00 | 0.15 | |
| ot1 | 9.65 | 1.00 | 0.00 | ** |
| ot2 | 1.52 | 1.00 | 0.22 | |
| ot3 | 1.31 | 1.00 | 0.25 | |
| Age | 0.89 | 1.00 | 0.35 | |
| RepetitionPairCount | 1.64 | 1.00 | 0.20 | |
| SES | 0.35 | 1.00 | 0.55 | |
| ot1:Age | 10.72 | 1.00 | 0.00 | ** |
| ot2:Age | 2.99 | 1.00 | 0.08 | . |
| ot3:Age | 2.21 | 1.00 | 0.14 | |
| ot1:RepetitionPairCount | 29.01 | 1.00 | 0.00 | *** |
| ot2:RepetitionPairCount | 0.16 | 1.00 | 0.69 | |
| ot3:RepetitionPairCount | 2.13 | 1.00 | 0.14 | |
| Age:RepetitionPairCount | 0.82 | 1.00 | 0.36 | |
| ot1:SES | 8.19 | 1.00 | 0.00 | ** |
| ot2:SES | 0.78 | 1.00 | 0.38 | |
| ot3:SES | 0.13 | 1.00 | 0.72 | |
| Age:SES | 0.55 | 1.00 | 0.46 | |
| RepetitionPairCount:SES | 0.31 | 1.00 | 0.57 | |
| ot1:Age:RepetitionPairCount | 21.70 | 1.00 | 0.00 | *** |
| ot2:Age:RepetitionPairCount | 0.05 | 1.00 | 0.83 | |
| ot3:Age:RepetitionPairCount | 3.43 | 1.00 | 0.06 | . |
| ot1:Age:SES | 9.82 | 1.00 | 0.00 | ** |
| ot2:Age:SES | 5.69 | 1.00 | 0.02 | * |
| ot3:Age:SES | 0.02 | 1.00 | 0.89 | |
| ot1:RepetitionPairCount:SES | 27.51 | 1.00 | 0.00 | *** |
| ot2:RepetitionPairCount:SES | 0.20 | 1.00 | 0.65 | |
| ot3:RepetitionPairCount:SES | 0.40 | 1.00 | 0.53 | |
| Age:RepetitionPairCount:SES | 0.03 | 1.00 | 0.86 | |
| ot1:Age:RepetitionPairCount:SES | 36.25 | 1.00 | 0.00 | *** |
| ot2:Age:RepetitionPairCount:SES | 5.27 | 1.00 | 0.02 | * |
| ot3:Age:RepetitionPairCount:SES | 0.72 | 1.00 | 0.40 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic), ot3 (cubic), Age in months, RepetitionPairCount denoting the repetition number of the image pair, and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 4.4:** Regression results for Reinforcement Trials at 18 months during Test

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 27.91 | 1.00 | 0.00 | *** |
| ot1 | 60.01 | 1.00 | 0.00 | *** |
| ot2 | 18.65 | 1.00 | 0.00 | *** |
| ot3 | 24.09 | 1.00 | 0.00 | *** |
| Age | 0.63 | 1.00 | 0.43 | |
| RepetitionPairCount | 13.45 | 1.00 | 0.00 | *** |
| SES | 23.46 | 1.00 | 0.00 | *** |
| ot1:Age | 25.03 | 1.00 | 0.00 | *** |
| ot2:Age | 0.14 | 1.00 | 0.71 | |
| ot3:Age | 2.95 | 1.00 | 0.09 | . |
| ot1:RepetitionPairCount | 218.66 | 1.00 | 0.00 | *** |
| ot2:RepetitionPairCount | 114.46 | 1.00 | 0.00 | *** |
| ot3:RepetitionPairCount | 97.69 | 1.00 | 0.00 | *** |
| Age:RepetitionPairCount | 6.48 | 1.00 | 0.01 | * |
| ot1:SES | 49.09 | 1.00 | 0.00 | *** |
| ot2:SES | 111.82 | 1.00 | 0.00 | *** |
| ot3:SES | 45.01 | 1.00 | 0.00 | *** |
| Age:SES | 16.59 | 1.00 | 0.00 | *** |
| RepetitionPairCount:SES | 27.12 | 1.00 | 0.00 | *** |
| ot1:Age:RepetitionPairCount | 108.61 | 1.00 | 0.00 | *** |
| ot2:Age:RepetitionPairCount | 8.28 | 1.00 | 0.00 | ** |
| ot3:Age:RepetitionPairCount | 24.48 | 1.00 | 0.00 | *** |
| ot1:Age:SES | 22.90 | 1.00 | 0.00 | *** |
| ot2:Age:SES | 96.78 | 1.00 | 0.00 | *** |
| ot3:Age:SES | 33.38 | 1.00 | 0.00 | *** |
| ot1:RepetitionPairCount:SES | 261.62 | 1.00 | 0.00 | *** |
| ot2:RepetitionPairCount:SES | 272.10 | 1.00 | 0.00 | *** |
| ot3:RepetitionPairCount:SES | 119.64 | 1.00 | 0.00 | *** |
| Age:RepetitionPairCount:SES | 21.57 | 1.00 | 0.00 | *** |
| ot1:Age:RepetitionPairCount:SES | 180.11 | 1.00 | 0.00 | *** |
| ot2:Age:RepetitionPairCount:SES | 253.93 | 1.00 | 0.00 | *** |
| ot3:Age:RepetitionPairCount:SES | 64.03 | 1.00 | 0.00 | *** |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic), ot3 (cubic), Age in months, RepetitionPairCount denoting the repetition number of the image pair, and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 4.5:** Regression results for Comprehension Trials at 18 months during Familiarisation

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 366.64 | 1.00 | 0.00 | *** |
| ot1 | 0.79 | 1.00 | 0.37 | |
| ot2 | 9.09 | 1.00 | 0.00 | ** |
| ot3 | 0.05 | 1.00 | 0.83 | |
| ot4 | 0.13 | 1.00 | 0.71 | |
| Age | 0.37 | 1.00 | 0.54 | |
| SES | 0.10 | 1.00 | 0.76 | |
| ot1:Age | 0.06 | 1.00 | 0.81 | |
| ot2:Age | 0.00 | 1.00 | 0.97 | |
| ot3:Age | 0.15 | 1.00 | 0.69 | |
| ot4:Age | 0.97 | 1.00 | 0.32 | |
| ot1:SES | 0.02 | 1.00 | 0.89 | |
| ot2:SES | 0.06 | 1.00 | 0.81 | |
| ot3:SES | 0.12 | 1.00 | 0.73 | |
| ot4:SES | 0.01 | 1.00 | 0.91 | |
| Age:SES | 0.16 | 1.00 | 0.69 | |
| ot1:Age:SES | 0.07 | 1.00 | 0.79 | |
| ot2:Age:SES | 0.08 | 1.00 | 0.77 | |
| ot3:Age:SES | 0.12 | 1.00 | 0.73 | |
| ot4:Age:SES | 0.12 | 1.00 | 0.73 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic), ot3 (cubic), ot4 (quartic), Age in months, and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 4.6:** Regression results for Comprehension Trials at 18 months during Test

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 171.54 | 1.00 | 0.00 | *** |
| ot1 | 2.99 | 1.00 | 0.08 | . |
| ot2 | 5.38 | 1.00 | 0.02 | * |
| ot3 | 1.37 | 1.00 | 0.24 | |
| ot4 | 0.79 | 1.00 | 0.38 | |
| Age | 1.16 | 1.00 | 0.28 | |
| SES | 0.13 | 1.00 | 0.72 | |
| ot1:Age | 1.63 | 1.00 | 0.20 | |
| ot2:Age | 0.02 | 1.00 | 0.90 | |
| ot3:Age | 0.30 | 1.00 | 0.59 | |
| ot4:Age | 0.15 | 1.00 | 0.70 | |
| ot1:SES | 0.91 | 1.00 | 0.34 | |
| ot2:SES | 0.00 | 1.00 | 0.95 | |
| ot3:SES | 0.23 | 1.00 | 0.63 | |
| ot4:SES | 0.41 | 1.00 | 0.52 | |
| Age:SES | 0.43 | 1.00 | 0.51 | |
| ot1:Age:SES | 1.34 | 1.00 | 0.25 | |
| ot2:Age:SES | 0.02 | 1.00 | 0.88 | |
| ot3:Age:SES | 0.45 | 1.00 | 0.50 | |
| ot4:Age:SES | 0.26 | 1.00 | 0.61 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic), ot3 (cubic), ot4 (quartic), Age in months, and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 4.7:** Regression results for Comprehension Trials split by Word Type at 18 months during Test

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 0.91 | 1.00 | 0.34 | |
| ot1 | 1.71 | 1.00 | 0.19 | |
| ot2 | 0.90 | 1.00 | 0.34 | |
| ot3 | 0.10 | 1.00 | 0.75 | |
| ot4 | 0.54 | 1.00 | 0.46 | |
| Age | 0.47 | 1.00 | 0.49 | |
| WordType | 142.11 | 2.00 | 0.00 | *** |
| SES | 0.00 | 1.00 | 0.96 | |
| ot1:Age | 1.22 | 1.00 | 0.27 | |
| ot2:Age | 0.89 | 1.00 | 0.35 | |
| ot3:Age | 0.01 | 1.00 | 0.90 | |
| ot4:Age | 0.84 | 1.00 | 0.36 | |
| ot1:WordType | 5.08 | 2.00 | 0.08 | . |
| ot2:WordType | 13.08 | 2.00 | 0.00 | ** |
| ot3:WordType | 0.27 | 2.00 | 0.88 | |
| ot4:WordType | 3.73 | 2.00 | 0.15 | |
| Age:WordType | 42.12 | 2.00 | 0.00 | *** |
| ot1:SES | 2.67 | 1.00 | 0.10 | |
| ot2:SES | 0.03 | 1.00 | 0.87 | |
| ot3:SES | 1.00 | 1.00 | 0.32 | |
| ot4:SES | 1.24 | 1.00 | 0.27 | |
| Age:SES | 0.15 | 1.00 | 0.69 | |
| WordType:SES | 1186.39 | 2.00 | 0.00 | *** |
| ot1:Age:WordType | 30.05 | 2.00 | 0.00 | *** |
| ot2:Age:WordType | 19.12 | 2.00 | 0.00 | *** |
| ot3:Age:WordType | 4.67 | 2.00 | 0.10 | . |
| ot4:Age:WordType | 0.71 | 2.00 | 0.70 | |
| ot1:Age:SES | 0.99 | 1.00 | 0.32 | |
| ot2:Age:SES | 0.06 | 1.00 | 0.81 | |
| ot3:Age:SES | 0.01 | 1.00 | 0.92 | |
| ot4:Age:SES | 0.01 | 1.00 | 0.94 | |
| ot1:WordType:SES | 300.23 | 2.00 | 0.00 | *** |
| ot2:WordType:SES | 64.90 | 2.00 | 0.00 | *** |
| ot3:WordType:SES | 5.65 | 2.00 | 0.06 | . |
| ot4:WordType:SES | 6.88 | 2.00 | 0.03 | * |
| Age:WordType:SES | 1576.01 | 2.00 | 0.00 | *** |
| ot1:Age:WordType:SES | 193.90 | 2.00 | 0.00 | *** |
| ot2:Age:WordType:SES | 75.14 | 2.00 | 0.00 | *** |
| ot3:Age:WordType:SES | 7.14 | 2.00 | 0.03 | * |
| ot4:Age:WordType:SES | 1.08 | 2.00 | 0.58 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic), ot3 (cubic), ot4 (quartic), Age in months, WordType including nouns, verbs ans adjectives, and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 4.8:** Regression results for Novelty Bias (Referent Selection Trials during Familiarisation) split by Familiar Image Type at 18 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 42.52 | 1.00 | 0.00 | *** |
| ot1 | 0.19 | 1.00 | 0.66 | |
| ot2 | 0.23 | 1.00 | 0.63 | |
| ot3 | 0.10 | 1.00 | 0.75 | |
| Age | 0.02 | 1.00 | 0.90 | |
| FamiliarityImage | 1.00 | 1.00 | 0.32 | |
| SES | 0.29 | 1.00 | 0.59 | |
| ot1:Age | 0.18 | 1.00 | 0.67 | |
| ot2:Age | 0.02 | 1.00 | 0.89 | |
| ot3:Age | 0.42 | 1.00 | 0.52 | |
| ot1:FamiliarityImage | 0.02 | 1.00 | 0.88 | |
| ot2:FamiliarityImage | 0.02 | 1.00 | 0.89 | |
| ot3:FamiliarityImage | 0.69 | 1.00 | 0.41 | |
| Age:FamiliarityImage | 0.98 | 1.00 | 0.32 | |
| ot1:SES | 0.05 | 1.00 | 0.83 | |
| ot2:SES | 0.17 | 1.00 | 0.68 | |
| ot3:SES | 1.51 | 1.00 | 0.22 | |
| Age:SES | 0.30 | 1.00 | 0.58 | |
| FamiliarityImage:SES | 0.08 | 1.00 | 0.78 | |
| ot1:Age:FamiliarityImage | 0.04 | 1.00 | 0.84 | |
| ot2:Age:FamiliarityImage | 0.02 | 1.00 | 0.90 | |
| ot3:Age:FamiliarityImage | 1.81 | 1.00 | 0.18 | |
| ot1:Age:SES | 0.07 | 1.00 | 0.79 | |
| ot2:Age:SES | 0.18 | 1.00 | 0.67 | |
| ot3:Age:SES | 1.09 | 1.00 | 0.30 | |
| ot1:FamiliarityImage:SES | 0.22 | 1.00 | 0.64 | |
| ot2:FamiliarityImage:SES | 0.11 | 1.00 | 0.74 | |
| ot3:FamiliarityImage:SES | 3.97 | 1.00 | 0.05 | * |
| Age:FamiliarityImage:SES | 0.15 | 1.00 | 0.70 | |
| ot1:Age:FamiliarityImage:SES | 0.24 | 1.00 | 0.63 | |
| ot2:Age:FamiliarityImage:SES | 0.09 | 1.00 | 0.76 | |
| ot3:Age:FamiliarityImage:SES | 4.26 | 1.00 | 0.04 | * |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months, FamiliarityImage indicates the type of familiar image (a familiar image that appeared in Reinforcement trials versus a familiar image that appeared in Comprehension trials), and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 4.9:** Regression results for Referent Selection Trials split by Word Type at 18 months during test

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 0.04 | 1.00 | 0.84 | |
| ot1 | 1.45 | 1.00 | 0.23 | |
| ot2 | 0.19 | 1.00 | 0.66 | |
| ot3 | 0.29 | 1.00 | 0.59 | |
| ot4 | 0.17 | 1.00 | 0.68 | |
| Age | 0.31 | 1.00 | 0.58 | |
| WordType | 2259.32 | 1.00 | 0.00 | *** |
| SES | 0.09 | 1.00 | 0.77 | |
| ot1:Age | 0.92 | 1.00 | 0.34 | |
| ot2:Age | 0.45 | 1.00 | 0.50 | |
| ot3:Age | 0.19 | 1.00 | 0.66 | |
| ot4:Age | 0.04 | 1.00 | 0.84 | |
| ot1:WordType | 3.44 | 1.00 | 0.06 | . |
| ot2:WordType | 2.63 | 1.00 | 0.10 | |
| ot3:WordType | 0.30 | 1.00 | 0.58 | |
| ot4:WordType | 37.15 | 1.00 | 0.00 | *** |
| Age:WordType | 1419.08 | 1.00 | 0.00 | *** |
| ot1:SES | 0.29 | 1.00 | 0.59 | |
| ot2:SES | 4.95 | 1.00 | 0.03 | * |
| ot3:SES | 0.26 | 1.00 | 0.61 | |
| ot4:SES | 0.18 | 1.00 | 0.68 | |
| Age:SES | 0.09 | 1.00 | 0.76 | |
| WordType:SES | 2235.40 | 1.00 | 0.00 | *** |
| ot1:Age:WordType | 36.94 | 1.00 | 0.00 | *** |
| ot2:Age:WordType | 7.83 | 1.00 | 0.01 | ** |
| ot3:Age:WordType | 0.64 | 1.00 | 0.42 | |
| ot4:Age:WordType | 43.00 | 1.00 | 0.00 | *** |
| ot1:Age:SES | 1.80 | 1.00 | 0.18 | |
| ot2:Age:SES | 5.05 | 1.00 | 0.03 | * |
| ot3:Age:SES | 0.88 | 1.00 | 0.35 | |
| ot4:Age:SES | 0.48 | 1.00 | 0.49 | |
| ot1:WordType:SES | 14.78 | 1.00 | 0.00 | *** |
| ot2:WordType:SES | 2.27 | 1.00 | 0.13 | |
| ot3:WordType:SES | 44.34 | 1.00 | 0.00 | *** |
| ot4:WordType:SES | 4.29 | 1.00 | 0.04 | * |
| Age:WordType:SES | 1839.66 | 1.00 | 0.00 | *** |
| ot1:Age:WordType:SES | 26.91 | 1.00 | 0.00 | *** |
| ot2:Age:WordType:SES | 0.97 | 1.00 | 0.32 | |
| ot3:Age:WordType:SES | 118.00 | 1.00 | 0.00 | *** |
| ot4:Age:WordType:SES | 0.05 | 1.00 | 0.82 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic), ot3 (cubic) and ot4 (quartic), Age in months, WordType indicates whether the target was the novel or the familiar, and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 4.10:** Regression results for Retention Trials at 18 months during familiarisation

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 0.00 | 1.00 | 0.99 | |
| ot1 | 0.76 | 1.00 | 0.38 | |
| ot2 | 0.03 | 1.00 | 0.86 | |
| ot3 | 0.31 | 1.00 | 0.58 | |
| Age | 0.01 | 1.00 | 0.92 | |
| SES | 0.22 | 1.00 | 0.64 | |
| ot1:Age | 0.27 | 1.00 | 0.60 | |
| ot2:Age | 0.28 | 1.00 | 0.60 | |
| ot3:Age | 0.23 | 1.00 | 0.63 | |
| ot1:SES | 8.98 | 1.00 | 0.00 | ** |
| ot2:SES | 1.35 | 1.00 | 0.24 | |
| ot3:SES | 0.37 | 1.00 | 0.54 | |
| Age:SES | 0.16 | 1.00 | 0.69 | |
| ot1:Age:SES | 6.94 | 1.00 | 0.01 | ** |
| ot2:Age:SES | 1.35 | 1.00 | 0.24 | |
| ot3:Age:SES | 0.08 | 1.00 | 0.78 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months, and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 4.11:** Regression results for Retention Trials at 18 months during test

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 6.71 | 1.00 | 0.01 | ** |
| ot1 | 0.50 | 1.00 | 0.48 | |
| ot2 | 0.09 | 1.00 | 0.76 | |
| ot3 | 0.00 | 1.00 | 0.97 | |
| Age | 9.20 | 1.00 | 0.00 | ** |
| SES | 31.72 | 1.00 | 0.00 | *** |
| ot1:Age | 0.32 | 1.00 | 0.57 | |
| ot2:Age | 0.10 | 1.00 | 0.75 | |
| ot3:Age | 0.27 | 1.00 | 0.60 | |
| ot1:SES | 0.07 | 1.00 | 0.79 | |
| ot2:SES | 1.16 | 1.00 | 0.28 | |
| ot3:SES | 0.02 | 1.00 | 0.89 | |
| Age:SES | 10.88 | 1.00 | 0.00 | *** |
| ot1:Age:SES | 0.20 | 1.00 | 0.66 | |
| ot2:Age:SES | 0.76 | 1.00 | 0.38 | |
| ot3:Age:SES | 0.07 | 1.00 | 0.79 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months, and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 4.12:** Regression results for ELP across all trials at 30 months during test

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 92.02 | 1.00 | 0.00 | *** |
| ot1 | 90.51 | 1.00 | 0.00 | *** |
| ot2 | 75.00 | 1.00 | 0.00 | *** |
| Age | 95.22 | 1.00 | 0.00 | *** |
| TrialType | 100.61 | 3.00 | 0.00 | *** |
| SES | 7.44 | 1.00 | 0.01 | ** |
| ot1:Age | 91.35 | 1.00 | 0.00 | *** |
| ot2:Age | 76.12 | 1.00 | 0.00 | *** |
| ot1:TrialType | 1195.54 | 3.00 | 0.00 | *** |
| ot2:TrialType | 337.99 | 3.00 | 0.00 | *** |
| Age:TrialType | 102.17 | 3.00 | 0.00 | *** |
| ot1:SES | 52.34 | 1.00 | 0.00 | *** |
| ot2:SES | 31.50 | 1.00 | 0.00 | *** |
| Age:SES | 6.41 | 1.00 | 0.01 | * |
| TrialType:SES | 8.60 | 3.00 | 0.04 | * |
| ot1:Age:TrialType | 1190.72 | 3.00 | 0.00 | *** |
| ot2:Age:TrialType | 338.44 | 3.00 | 0.00 | *** |
| ot1:Age:SES | 54.39 | 1.00 | 0.00 | *** |
| ot2:Age:SES | 32.83 | 1.00 | 0.00 | *** |
| ot1:TrialType:SES | 197.72 | 3.00 | 0.00 | *** |
| ot2:TrialType:SES | 52.19 | 3.00 | 0.00 | *** |
| Age:TrialType:SES | 7.43 | 3.00 | 0.06 | . |
| ot1:Age:TrialType:SES | 207.09 | 3.00 | 0.00 | *** |
| ot2:Age:TrialType:SES | 52.99 | 3.00 | 0.00 | *** |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear) and ot2 (quadratic), Age in months, and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 4.13:** Regression results for Reinforcement Trials at 18 and 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 5.83 | 1.00 | 0.02 | * |
| ot1 | 5.85 | 1.00 | 0.02 | * |
| ot2 | 0.01 | 1.00 | 0.94 | |
| Age | 6.76 | 1.00 | 0.01 | ** |
| RepetitionPairCount | 17.76 | 1.00 | 0.00 | *** |
| TestType | 0.48 | 1.00 | 0.49 | |
| SES | 21.77 | 1.00 | 0.00 | *** |
| ot1:Age | 6.39 | 1.00 | 0.01 | * |
| ot2:Age | 0.02 | 1.00 | 0.89 | |
| ot1:RepetitionPairCount | 2.81 | 1.00 | 0.09 | . |
| ot2:RepetitionPairCount | 0.85 | 1.00 | 0.36 | |
| Age:RepetitionPairCount | 18.48 | 1.00 | 0.00 | *** |
| ot1:TestType | 6.75 | 1.00 | 0.01 | ** |
| ot2:TestType | 0.37 | 1.00 | 0.54 | |
| Age:TestType | 1.11 | 1.00 | 0.29 | |
| RepetitionPairCount:TestType | 12.11 | 1.00 | 0.00 | *** |
| ot1:SES | 0.00 | 1.00 | 0.97 | |
| ot2:SES | 0.97 | 1.00 | 0.32 | |
| Age:SES | 21.84 | 1.00 | 0.00 | *** |
| RepetitionPairCount:SES | 55.58 | 1.00 | 0.00 | *** |
| TestType:SES | 2.24 | 1.00 | 0.13 | |
| ot1:Age:RepetitionPairCount | 2.71 | 1.00 | 0.10 | . |
| ot2:Age:RepetitionPairCount | 0.82 | 1.00 | 0.36 | |
| ot1:Age:TestType | 8.30 | 1.00 | 0.00 | ** |
| ot2:Age:TestType | 0.34 | 1.00 | 0.56 | |
| ot1:RepetitionPairCount:TestType | 60.81 | 1.00 | 0.00 | *** |
| ot2:RepetitionPairCount:TestType | 0.32 | 1.00 | 0.57 | |
| Age:RepetitionPairCount:TestType | 14.52 | 1.00 | 0.00 | *** |
| ot1:Age:SES | 0.04 | 1.00 | 0.84 | |
| ot2:Age:SES | 1.05 | 1.00 | 0.31 | |
| ot1:RepetitionPairCount:SES | 1.21 | 1.00 | 0.27 | |
| ot2:RepetitionPairCount:SES | 0.03 | 1.00 | 0.87 | |
| Age:RepetitionPairCount:SES | 55.63 | 1.00 | 0.00 | *** |
| ot1:TestType:SES | 2.77 | 1.00 | 0.10 | . |
| ot2:TestType:SES | 0.70 | 1.00 | 0.40 | |
| Age:TestType:SES | 4.87 | 1.00 | 0.03 | * |
| RepetitionPairCount:TestType:SES | 30.80 | 1.00 | 0.00 | *** |
| ot1:Age:RepetitionPairCount:TestType | 31.21 | 1.00 | 0.00 | *** |
| ot2:Age:RepetitionPairCount:TestType | 0.02 | 1.00 | 0.89 | |
| ot1:Age:RepetitionPairCount:SES | 1.02 | 1.00 | 0.31 | |
| ot2:Age:RepetitionPairCount:SES | 0.01 | 1.00 | 0.93 | |
| ot1:Age:TestType:SES | 1.93 | 1.00 | 0.17 | |
| ot2:Age:TestType:SES | 0.25 | 1.00 | 0.61 | |
| ot1:RepetitionPairCount:TestType:SES | 56.90 | 1.00 | 0.00 | *** |
| ot2:RepetitionPairCount:TestType:SES | 21.17 | 1.00 | 0.00 | *** |
| Age:RepetitionPairCount:TestType:SES | 41.58 | 1.00 | 0.00 | *** |
| ot1:Age:RepetitionPairCount:TestType:SES | 26.11 | 1.00 | 0.00 | *** |
| ot2:Age:RepetitionPairCount:TestType:SES | 8.11 | 1.00 | 0.00 | ** |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear) and ot2 (quadratic), Age in months, and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 4.14:** Regression results for Comprehension Trials at 18 and 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 0.00 | 1.00 | 0.96 | |
| ot1 | 0.00 | 1.00 | 0.96 | |
| ot2 | 0.01 | 1.00 | 0.91 | |
| ot3 | 0.05 | 1.00 | 0.82 | |
| TestType | 1.27 | 1.00 | 0.26 | |
| Age | 0.00 | 1.00 | 0.95 | |
| SES | 0.00 | 1.00 | 0.99 | |
| ot1:TestType | 0.99 | 1.00 | 0.32 | |
| ot2:TestType | 0.94 | 1.00 | 0.33 | |
| ot3:TestType | 0.01 | 1.00 | 0.91 | |
| ot1:Age | 0.00 | 1.00 | 0.98 | |
| ot2:Age | 0.01 | 1.00 | 0.93 | |
| ot3:Age | 0.01 | 1.00 | 0.93 | |
| TestType:Age | 1.11 | 1.00 | 0.29 | |
| ot1:SES | 0.01 | 1.00 | 0.93 | |
| ot2:SES | 0.01 | 1.00 | 0.90 | |
| ot3:SES | 0.30 | 1.00 | 0.59 | |
| TestType:SES | 1.10 | 1.00 | 0.29 | |
| Age:SES | 0.00 | 1.00 | 1.00 | |
| ot1:TestType:Age | 0.84 | 1.00 | 0.36 | |
| ot2:TestType:Age | 0.82 | 1.00 | 0.37 | |
| ot3:TestType:Age | 0.01 | 1.00 | 0.92 | |
| ot1:TestType:SES | 0.84 | 1.00 | 0.36 | |
| ot2:TestType:SES | 1.28 | 1.00 | 0.26 | |
| ot3:TestType:SES | 0.22 | 1.00 | 0.64 | |
| ot1:Age:SES | 0.00 | 1.00 | 0.96 | |
| ot2:Age:SES | 0.01 | 1.00 | 0.92 | |
| ot3:Age:SES | 0.23 | 1.00 | 0.63 | |
| TestType:Age:SES | 0.94 | 1.00 | 0.33 | |
| ot1:TestType:Age:SES | 0.71 | 1.00 | 0.40 | |
| ot2:TestType:Age:SES | 1.12 | 1.00 | 0.29 | |
| ot3:TestType:Age:SES | 0.25 | 1.00 | 0.62 | |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months, and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

# Chapter 5

# Language Input and Early Language Processes in India

## 5.1 Introduction

Extensive work in western cultures has shown multiple relationships between early language input and later language abilities (e.g., Rowe, 2012), as well as the role that contextual variables play in language development (e.g., Hoff, 2006). However, the role of language input in typical language development is less clear once we take a broad view of the world's language learning environments. This is because most of our data comes from children living in western industrialised areas, and mostly from children learning English. Other languages and cultures around the word are underrepresented and we do not know how other children's experiences shape their paths towards developing the necessary language skills to successfully communicate in their society. The present project aims to fill this gap by presenting early language input data in relation to later language processing abilities in children living in a rural village, Shivgarh, located in Uttar Pradesh, India.

### 5.1.1 Language Input, Outcomes and SES in the West

Within Western-influenced industrialized cultures, the amount and type of speech addressed to infants and young children has a big impact on their language development (Hart & Risley, 1995; Hoff-Ginsberg, 1991; Hurtado et al., 2008b; Huttenlocher et al., 2002). These effects have been found across a variety of situations such as in-lab experiments, play sessions and naturalistic home recordings. For example, Hurtado et al. (2008a) examined maternal speech at 18 months in relation to children's speech processing efficiency and vocabulary at 18 and 24 months on Spanish learning children. Children with mothers who provided larger quantities of language input at 18 months knew more words and were faster in word recognition at 24 months. Similarly, Weisleder and Fernald (2013) found that 19-month-old infants who experienced more child-directed speech were also more efficient at processing familiar spoken words. Moreover, those children had larger expressive vocabularies at 24 months. Similar evidence has been found in older children. Mahr and Edwards (2018) used looking responses in preschool age children to measure how lexical processing as well as language input at 28–39 months predicted vocabulary size (assessed using direct measures) one year later. The authors found that language input and lexical processing predicted receptive vocabulary growth, indicating that both language experience as well as processing abilities are related to vocabulary development. Thus, it is possible that larger quantities of input provided children more opportunities to practice recognizing words, which led to greater processing efficiency, facilitating word learning. These studies provide examples of how language input relates to later language ability as a function of the child's linguistic and cognitive capacities, particularly speed of word processing.

A source of variation in children's language exposure in Western contexts appears linked to the social class (or socio-economic status, SES) of the child's family often computed based on the income and/or parental education (e.g., Hoff, 2003). It is well established that children from lower SES families build their vocabularies at slower rates than children from higher SES families (e.g., Dollaghan et al., 1999; Hart & Risley, 1995; Rowe, 2008). For example, in a study with middle-class families, Huttenlocher et al. (1991) found that only the amount of parental input predicted vocabulary growth between 14 and 26 months. In their seminal study, (Hart & Risley, 1995) looked at variation in the quantity of input across the early childhood period in families ranging in SES. They estimated that by the time children reach school age, those growing up in higher-SES families were, on average, exposed to 30 million more words than children growing up in lower-SES families. However, this idea has been recently debated with the counter argument that parental linguistic input may be a limited indicator for certain groups such as lower SES families (see Sperry et al., 2019, 2019 for studies showing that low SES children might hear more speech that previously reported, and Golinkoff et al., 2019 for a response acknowledging the existence of the 30-million-word gap). Hart and Risley (1992, 1995) also found strong positive associations between quantity of caregiver input and children's vocabulary growth, supporting the notion that the quantity of parental vocabulary input influences children's rate of vocabulary growth. Quality of input also mattered with higher-SES parents responding more to their children, producing more affirmative and encouraging instances and fewer prohibitions. In addition, high SES parents showed more diverse input because they produced more noun types and modifiers per hour. Similar findings were reported in a Family Life study with a large homogeneous sample of 1,292 children followed from birth (Vernon-Feagans et al., 2020). The authors found that maternal language input, indexed as the number of different words, mean length of utterance and number of wh-questions, partially mediated the relationship between maternal education and later child language at school age. Thus, both quantity and quality of language input contribute to later language skill but in different ways.

A study by Hoff and Naigles (2002) showed that 2-year-old children's lexical development benefited from higher quantity, more lexical richness, and more syntactic complexity of maternal language input in data from mother-child conversation during dyadic play. Benefits from both quantity and quality of input have also been found across SES groups and beyond vocabulary skills. Huttenlocher et al. (2010) followed a group of children with diverse SES backgrounds longitudinally from 14 to 46 months to examine the role of quantity of input (e.g., word tokens) and diversity of input (i.e., variety of words and syntactic structures) in children's vocabulary and syntactic growth. In this study, variations in language input, particularly differences in the syntactic structures caregivers used, affected children's language growth. Further, while quantity and diversity of input was related to SES, diversity of caregiver speech was a significant predictor of child vocabulary growth, measured as the word types children produced when controlling for SES.

This body of research highlights the role of SES on language input and language development, but, what makes SES such a powerful predictor of later language skills? An aim of current research is to identify the pathways by which SES influences children's language skills. This is not an easy task because SES and child development are multifaceted variables and many factors that influence child development are correlated with SES. This makes it difficult to identify the causal relations underlying SES effects on child development (Hoff et al., 2002). Nevertheless, it has been proposed that maternal speech mediates the relation between SES and child vocabulary development

(Huttenlocher et al., 1991; Hoff et al., 2002; Hoff, 2003). In a study looking at high and low SES mother-child dyads in North America, 2-year-old children from higher SES families had larger productive vocabularies than their low SES peers; critically, these differences were fully explained by properties of maternal speech that differed as a function of SES (Hoff, 2003). Interestingly, proprieties of maternal speech captured both linguistic properties of the input the children received (e.g., number of types, tokens and utterances) and social properties of the interactions they experienced (e.g., number of utterances in episodes of joint attention or topic-continuing replies) which builds on evidence showing that both quantity and quality of language input matter for language development. It is clear that, in western cultures, the role of maternal speech is crucial for children's language abilities, and thus, variations in mother's speech based on her education lead to variations in input. Maternal speech might be highly influential in cultures where mothers are the primary caregiver, because children likely hear most of their input from mothers. The question is then whether this translates to other cultures and societies where this is not the case.

### 5.1.2 Language Input and Outcomes in other Cultures

Studies have documented large qualitative differences across cultures in the ways in which infants are brought into the social and cultural world they grow up in (e.g., Brown & Gaskins, 2014). Differences in children's socio-cultural context also have an effect on children's language experiences. Factors such larger families, social constructs regarding parenting and language development, cultural practices, and daily routines shape the way a community interacts with children, and therefore, the way they speak with them and/or around them. For example, in western industrialised cultures, infants are most often raised in nuclear families with few siblings and with primary care being given by one or two main caregivers. In non-western societies and indigenous communities, however, children are often raised in large families and different caregivers might be involved in their upbringing. Thus, we could imagine that children's early experiences in western contexts are usually highly influenced by those primary caregivers, whereas in non-western contexts children might be more dependent on their language experiences with the whole family. Even though there are large variations in parenting styles across different cultures, normally developing children around the world learn language. This means that it is possible that children are able to adapt and use the cues available in their environment in different ways, to build their language skills based on their unique language experiences.

This can be illustrated with studies in non-western cultures looking at the effect of child directed speech (CDS) in communities were CDS is very rare. In western populations, child directed speech has been reported to facilitate early word learning (Cartmill et al., 2013; Hoff, 2003; Rowe, 2008; Weisleder & Fernald, 2013). However, studies in rural or Indigenous populations report that children hear less child directed speech (Vogt & Mastin, 2013; Shneidman & Goldin-Meadow, 2012; Casillas et al., 2020; Cristia et al., 2019) than children in urban and western settings (e.g., children from United States and Canada Bergelson et al., 2019). In some communities, children are mostly exposed to overheard speech (see Casillas et al., 2020 for language input data gathered in a Tseltal Mayan village and Cristia et al., 2019 for data gathered on the Tsimane community), or speech from other children (Shneidman & Goldin-Meadow, 2012). Still, children in those studies do not show an apparent delay on indicators of children's language milestones (e.g., babbling, first words, first word combinations) and they grow up to become competent users of the language spoken around them (see Lieven, 1994). In fact, no differences have been found either in the emergence of communicative behaviours (e.g., pointing) across very different cultures (see Lieven & Stoll, 2013 for a comparison between Germany and Nepal).

Nevertheless, in a recent study by Shneidman and Goldin-Meadow (2012), the authors investigated naturally occurring language input to Yucatec Mayan children. They compared this input to the input heard by children growing up in large families in the United States (2 - 4 family members including other children), and measured how directed and overheard input related to Mayan children's later vocabulary. Their findings showed that 1-year-old Mayan children heard a smaller proportion of total input in directed speech than children from the US. Also, they found that for Mayan (but not US) children, there was a greater increase in the proportion of directed input that children receive between 13 and 35 months, with a large part of this input coming from other children. Finally, the study showed that the number of word types directed to Mayan children from adults at 24 months (but not word types overheard by children or word types directed from other children) predicted later vocabulary (calculated as a composite score between productive and receptive vocabulary). These findings suggest that adult talk directed to children is important for early word learning, even in communities where much of children's early language input comes from overheard speech.

Another study on an ethnic group where infants are rarely spoken to (the Tsimane, based in Bolivia) has documented the long term impact of low amounts of language input in childhood on phonological processing in children and adults. In this study, the authors found lower non-word repetition (NWR) scores than in previous work for both children and adults, which is consistent with the hypothesis that there would be long-term effects on phonological processing from experiencing low levels of directed input in infancy (Cristia, Farabolini, et al., 2020). Furthermore, recent evidence shows that even though children from communities with low CDS reach expected basic linguistic milestones, Tseltal and Yélî children's early consonant production demonstrates greater environmental sensitivity than canonical babble. This suggests that vocal maturity measures might be fairly robust to environmental variation, whereas consonant acquisition might be environmentally sensitive. This result may be in line with previous literature showing that variation in child directed speech is associated with variation in lexical development (Peute & Casillas, 2021).

These data together illustrate how different early language experiences can be across children growing up around the world. They also highlight the relevance of studying language input and its relation to later language abilities in different contexts, thereby informing theories of language development. These studies also reveal that children are highly adaptable and that they might use different strategies to learn from their input across very different environments. Some studies have started looking at the relationships between language input and language abilities in non-western societies. However, those studies are scarce and they do not document how input relates to the specific language processes involved in word learning abilities, which is crucial to better understand how children develop language based on to their linguistic experiences. Measuring language abilities in young infants is challenging in any society. There are not many direct measures of language abilities early in development and even fewer that relate different processes involved in word learning to each other. This requires culturally relevant tasks that can be easily administered in any cultural setting. Moreover, studies show that adult-directed speech has been related to vocabulary outcomes even in cases when this is rare. This means that children manage to extract the linguistic information they need despite minimal directed speech. Yet, it is not clear how variation in the amount of input children are exposed to relates to later language processing abilities. It is also not clear whether the impact of SES on language input and language abilities generalizes beyond non-western cultures. Thus, further research is needed to understand how language input relates to language outcomes, and the role that environmental variables might play across different cultures and populations.

The present project aims to contribute to this literature by investigating how early language input relates to later language processing abilities as a functions of maternal education, in children living in a rural village, Shivgarh, located in Uttar Pradesh, India. To understand how language input influences developing language processes at different time points, we used longitudinal data relating early children's language environment to later linguistic processing abilities. Children's home language input was measured using the LENA system; language processes were measured using the ELP task (developed in Chapter 2), a direct, culturally-valid measure of children's language processing abilities that was adapted to the language and culture of the Indian sample (see Chapter 4 for details regarding this adaptation). The ELP task is an eye-tracking based task that is able to capture multiple language processes of individual children by combining different measures of language processing in a relatively short assessment lasting 15 - 20 minutes. This allows the researcher to examine how constellations of language processes influence one another. The ELP task measures multiple processes shown to be predictive of later language skill: speed of word processing based on work from A. Fernald et al. (1998), novelty bias, referent selection (or disambiguation), and retention of new words based on studies using the Reference Selection and Retention (RSR) task with either 2D or 3D images (Bion et al., 2013; Samuelson et al., 2017; Horst & Samuelson, 2008). ELP also incorporates an online measure of word comprehension which gives a direct measure of a child's vocabulary size or word comprehension abilities, which is based on the Computerised Comprehension Task (CCT) (Friend & Keplinger, 2008). In the western context, language input, language outcomes and processing abilities have been associated as important features for language learning. For instance, children who are exposed to more words have larger vocabularies and faster word processing abilities which also helps them increase their vocabulary skills. These relationships have not be explored in non-western samples.

The current study was conducted in Shivgarh, a remote village located in rural Uttar Pradesh, India. The state of Uttar Pradesh is the most populous state in India and scores amongst the worst regions in terms of human development indicators. Due to the socio-cultural, demographic and health system characteristics, Uttar Pradesh accounts for a quarter of India's neonatal deaths and for 8% of neonatal deaths worldwide. This is why this regions has been a focus of targeted interventions to decrease neonatal mortality (Kumar et al., 2008). In rural Uttar Pradesh, in 2017 - 2018, the literacy rate among men was 80.5% and among women was 60.4%, this included persons aged 7 and above (data extracted from a report on education in India as part of 75th round of National Sample Survey *Key Indicators of Household Social Consumption on Education in India. NSS 75th Round*, 2018). The local dialect is called Awadhi, and it was reported to have 4 million native speakers in 2011 in India (*Language. India, States and Union Territories*, 2011). Awadhi belongs to the Indo-European languages, particularly the Indo-Aryan sub-family. It is generally viewed as a rural tongue, yet people in urban areas tend to speak a mixed form of Awadhi with standard Hindi, whereas in rural areas people speak only Awadhi. Education in rural areas in Uttar Pradesh is in Hindi but there is considerable epic literature written in Awadhi (for a detailed study of Awadhi language and its characteristics see Saksena, 1971). Indian families from Shivgarh are very large, as is often the case in rural areas in India (*Language. India, States and Union Territories*, 2011). When a baby is born, mothers are the primary carers of the newborn but basic practices diverge based on socio-demographic factors (see Baqui, 2007 for a medical description of the newborn care in this area and the effects of demographic factors such as maternal education and Darmstadt et al., 2008, for a description of child care in Shivgarh as well as community perceptions of birth weight). This leads in several cases to malnutrition and stunting in infancy (Brennan, McDonald, & Shlomowitz, 2004). At older ages, parents are usually involved in farming work during the day and grandparents as well as older siblings help to take care for their younger children in the families (see Sahithya, Manohari, & Vijaya, 2019). This features make children from Shivgarh an interesting case to measure children's early language experiences in relation to later skills.

To date, language development studies in non-western samples have usually focused on rural areas or indigenous communities, but none to our knowledge has looked at communities with children at-risk due to poverty. Thus, in this project, we are not only measuring the relationship between early input and language outcomes, but also the impact of maternal education and poverty. Poverty has been associated with cognitive development, including language skill. Children experiencing early adversity, such as nutritional deficits are at a high risk of delays in their cognitive development (Leroy et al., 2012; L. C. Fernald et al., 2012). This represents a high percentage of children in low and middle income countries. It is estimated that 250 million children (about a 43%) in low and middle income countries fail to reach their developmental potential due to early adversity (Black et al., 2017). Poverty and early adversities significantly impact development, accentuating the risk of poor socioeconomic outcomes and contributing to a cycle of poverty. A challenge is that studies looking at early adversity document SES effects across low and middle income countries on children's cognitive abilities (e.g., L. C. Fernald et al., 2012), but they do not measure the effect of adversity on specific mechanisms that support early word learning. Thus, we do not know how findings generalise to other populations. Associations between socioeconomic status (SES) and early growth faltering (i.e., stunting), memory, executive function (L. C. Fernald et al., 2012), brain development (Hackman & Farah, 2009; Noble et al., 2012) and language (Hart & Risley, 1995) have been well established in infancy and childhood and across the lifespan (Bradley & Corwyn, 2002; Brooks-Gunn & Duncan, 1997; Kelly et al., 2011).

### 5.1.3 The Present Study

To have a better picture of how the processes that support language development are impacted by language input, the present study investigated the relationship between home language input at two different time points, infancy and toddlerhood, and early language processes in a sample of Indian infants. Adult input and conversational turns were measured in children's natural environment using the LENA system. Subsequently, children's language processing abilities including speed of word processing, comprehension abilities, novelty biases, referent selection and retention of new words were measured using the Early Language Processing task.

Our daylong LENA recordings lasted up to 16 hours. Day-long recordings may yield more representative data on children's language experiences than shorter samples of language input (e.g., gathered in the lab during parent-child interactions), because they are more likely to capture language in different contexts that are more representative of children's day-to-day lives. In a recent meta-analysis by Anderson, Graham, Prime, Jenkins, & Madigan, 2021, the length of the observation significantly moderated the association between the quality of parental linguistic input and child language, with longer observation periods leading to larger effect sizes.

To measure language processing abilities, we tested children at 18 months-of-age using the ELP task. Chapter 4 details how we developed this portable, efficient eye-tracking based task and adapted it to another language and culture. Finally, we examined the association between demographic variables and language input as well as language outcomes in the rural India sample. We found evidence of SES effects, based on maternal education, on the ELP data we report on Chapter 4 for this population. Thus, an additional aim of this project is to measure the impact of environmental variables such as SES, indexed via maternal education, on children's language input and its subsequent relationships to later language processes. In this context, SES is also an index of poverty, since children from lower SES in our sample were more likely to be exposed to a higher degree of adversity than their peers (e.g., malnutrition, low weight or poorer living conditions). This study extends the findings of Chapter 3 to an Indian population.

## 5.2 Methods

The goal of this study is to asses how children's language input in infancy and toddlerhood measured with LENA influenced their language processing abilities at 18 months measured with the ELP task. Towards this aim, we collected LENA data at two different time points and ELP data at 18 months from the same group of children. Because we were also interested in the possible effects of children's environment, we also collected contextual information such as the family SES.

### 5.2.1 Participants

The main analysis of this chapter includes children from which we have two LENA observations as well as ELP data at 18 months. Thus, the data is a sub-sample of the ELP data reported in Chapter 4.

Demographic information for the subsample of 82 children with with two LENA observations (35 female) can be seen in Table 5.1. Children in our sample included 40 from High SES families and 42 from Low SES families. In this study, SES is calculated based on maternal education (low = illiterate or primary education; high = greater than middle school). The terms low and high are relative to our sample.

The first LENA measurement, "LENA Infant," was collected when the children were between 4 and 13 months of age ($M$ = 9.87 months, $SD$ = 1.61 months). The second LENA measurement, "LENA Toddler," was collected when the children were between 14 and 25 months of age ($M$ = 16.94 months, $SD$ = 2.32 months). From the initial sample of 108 children with LENA recordings, 24 did not have data for both observations (Infant and Toddler) and 2 were excluded because there was no data for one or more of the LENA measures, indicating that the recording was likely to have failed.

The final sample of children with LENA Infant, LENA Toddler and ELP data at 18 months (range 17 - 22 months) included a total of 46 children (19 girls; 26 Low SES).

This project was reviewed and approved by the CEL ethics committee. Parents signed or provided an oral videotaped informed consent form (in cases

| **LENA Sample Demographics For India**<br>Participants with two observations N = 82 (35 girls) | |
|---|---|
| **Age in Months** | |
| **LENA Infant** | |
| Mean (SD) | 9.9 (1.6) |
| Median [Min, Max] | 10 [4.3, 13] |
| **LENA Toddler** | |
| Mean (SD) | 16.9 (2.3) |
| Median [Min, Max] | 16.8 [13.6, 25.7] |
| **SES** | |
| Low | 42 (51.2%) |
| High | 40 (48.8%) |
| **Electricity** | |
| Yes | 48 (58.5%) |
| No | 34 (41.5%) |
| **Caste** | |
| Scheduled caste-scheduled tribe (traditionally most depressed) | 58 (70.7%) |
| Other backwards caste (socially or economically disadvantaged) | 21 (25.6%) |
| General (middle class) | 3 (3.7%) |
| **Family members** | |
| 3-4 | 19 (23.2%) |
| 5-8 | 48 (58.5%) |
| more than 8 | 15 (18.3%) |
| **Mother's Education Status** | |
| Illiterate/non-primary | 22 (26.8%) |
| Primary school | 20 (24.4%) |
| High school | 20 (24.4%) |
| Some higher education | 20 (24.4%) |
| **Income PPP** | |
| $\leq$ 2000 INR | 32 (39.0%) |
| between 2000 and 4000 INR | 28 (34.1%) |
| > 4000 INR | 22 (26.8%) |

**Table 5.1:** Demographic Information of LENA data in India.

where the caregiver was illiterate). The subset of the data reported here is part of a larger study examining the early precursors of executive function in India led by Prof. John Spencer. The project was funded by a grant of the Bill & Melinda Gates Foundation (No. OPP1164153) to Prof. John Spencer.

### 5.2.2 Procedure

At each LENA time point (LENA Infant and LENA Toddler), members of the CEL team went to each participant home and gave them a LENA audio recording device after instructing them how to use it. After 16 hours (the maximum recording length of a LENA device), CEL members collected the device. In total we gathered 5021.09 hours of recording data.

At approximately 18 months, the same children came to the CEL Laboratory for an ELP session. The procedure is reported in Chapter 4. Children sat on their parent's lap. An Eye-Link Duo (SR Research, Ontario, Canada) eye-tracker in the remote setting captured children's gaze. Since these families were part of a larger longitudinal project, the same day that they did the ELP task, most children also participated in a dyadic free-play session and a visual working memory task involving eye-tracking as well. Children had breaks in between sessions or came another day if they got tired of participating in the different studies.

### 5.2.3 Data Processing and Analytical Approach

**LENA data:** The home audio recordings were exported using the LENA proprietary software. The advanced data extraction software (ADEX) from LENA provided several estimates of the child's language environment, including Adult Word Count (AWC), defined as the number of words spoken in the vicinity of the child, Child Vocalization Count (CVC), defined as the number of vocalizations (including words and non-words, such as babbling or exclamations such as *ah!*), the child's exposure to non-social electronic media (e.g., TV, radio, music), and child–adult conversational turn count (CTC), defined as two discrete utterances between child–adult pairs that contain a pause of no longer the 5 seconds. Note that CTC is a composite measure that contains AWC and CVC that happened consecutively. We used the Vocalisation Activity Block (One Row per Block per Recording) including Segment details as output, because it gave the highest resolution. From each extended home recording at each LENA time point we found the hour with highest AWC, CTC and CVC by first extracting counts for 1 hour bins across the entire recording (as in Romeo et al., 2018). The maximum count per each LENA measure was then selected across the different LENA days. This gave a maximum AWC, CTC and CVC per hour for each LENA observation. This processing was done using the statistical package R (R Core Team, 2017).

We had three main questions regarding the language measures extracted from LENA (AWC, CTC and CVC): Q1) Do AWC, CTC and CVC differ across the two time points? Q2) Were AWC, CTC and CVC consistent across individuals at both time points? Q3) Were LENA measures related to SES differences in our sample?

To answer Q1 we ran a Wilcoxon test (a more conservative version of a t-test) to measure if there were differences between LENA measures at both time points. To answer Q2, we ran correlation analyses to look at possible relationships between LENA measures during infancy and LENA measures during toddlerhood. To answer Q3, we ran linear models to assess possible relationships between LENA and SES status (i.e., SES based on maternal education). In this set of models, LENA Infant and SES predicted LENA Toddler measures. All LENA measures and SES were centered. Model fit was assessed using the *check_model* function from the R package Performance (Lüdecke et al., 2021), which generates a visual check of various model assumptions such as normality of residuals, normality of random effects, linear relationship, homogeneity of variance and multicollinearity.

In all of our LENA data analyses, we modeled the three LENA outcome measures separately (AWC, CTC and CVC), as a quality check. For example, one would expect an increase in child vocalisations as the child ages as well as a strong correlation between child vocalisations and conversational turn counts.

**ELP data:** The eye-tracking data from ELP were pre-processed using Data Viewer (SR-Research, Ontario, Canada). Trials were segmented into periods of interest (IP) using message-based events. Areas of interest (AOI) were set to be 50% bigger than target objects to account for calibration errors and drifts in the eye tracker. Sample reports were exported and raw gaze position was processed using the statistical package R (R Core Team, 2017).

Eye-tracking data from word onset to 1800 ms after onset from the ELP test phase was processed using the eyetrackingR package (Dink & Ferguson, 2016). During data processing, trials with more than 40% of trackloss were removed from the analysis. Mean proportions of looks to the target per each ELP trial type (Reinforcement, Comprehension, Referent Selection and Retention), as well as proportions of looks to the novel item during familiarisation of Referent Selection trials (for Novelty Bias) were extracted following the same procedure as in Chapter 4.

The main question was whether language input at two time points, early in infancy and later in toddlerhood, is associated with language processes at 18 months. To examine this, we ran a set of linear models with LENA measures at both time points as well as SES predicting ELP measures at 18 months. The main analysis examining LENA and ELP relationships did not look at child vocalisations because our focus was on the linguistic input (AWC and CTC) rather than on the child's productions, CVC.

LENA input measures were included in each model as fixed terms and entered separately because they are highly correlated due to the fact that turn count is calculated using adult words and child vocalisations. Thus, we never modeled AWC and CTC together. Both LENA input measures as well as SES were scaled and centered. The predicted ELP measures were mean proportion of looking to target in ELP Reinforcement, Comprehension, Referent Selection, and Retention trails, each modeled separately. We also included ELP Novelty Bias as an additional predicted measure. In addition to the LENA measures as part of the fixed effects, models predicting ELP Reinforcement included SES (High SES or Low SES) as fixed effects. Models predicting ELP Comprehension included Word Type (noun, verb and adjective) and SES (High SES or Low SES) as fixed effects. These models also included a random effects structure where a random intercept was nested within participant; this allowed a random slope to be fitted to each participant. Models predicting ELP Novelty Bias, included Familiar Image Type and SES (High SES or Low SES) as predictors. Familiar Image Type refers to the type of familiar image that was paired with the novel one. This image could be familiar or highly familiar. Familiar images also appeared in ELP Comprehension, highly familiar images also appeared on ELP Reinforcement. Models predicting ELP Referent Selection also included Word Type, which refers to whether the target was novel or a familiar noun as predictors as well as SES as fixed effects. Finally, models predicting ELP Retention included SES (High SES versus Low SES) in the fixed effects structure in addition to the LENA observations.

Model fit to the data for all was assessed using the DHARMa R package (Hartig, 2021), which uses a simulation-based approach to create readily interpretable scaled (quantile) residuals for fitted (generalized) linear mixed models, in addition the Performance R package (Lüdecke et al., 2021). For each model, the effect of each parameter was assessed with an F test, in particular, we used the ANOVA function from the car R package (R Core Team, 2017), which tests whether the model terms are significant. All the reported effects and interactions are those that remained after using this method.

## 5.3 Results

The main aim of this study was to examine whether language input predicts children's language processing abilities while taking contextual variables such as SES into account. We first report analysis on the LENA data only, characterizing the amount of input children heard, their own turn and vocalisation productions, and changes in these variables over time. Then we examine whether the language input the children heard as infants and toddlers predicted language processing measures at 18 months.

### 5.3.1 Language Input

The first question was whether the three LENA measures of adult input (AWC), conversational turns (CTC) and child vocalisations (CVC) differed across LENA observations (Infant versus Toddler). As can be seen in Figure 5.1, paired Wilcoxon signed-rank tests showed that the number of adult words, indexed by the maximum adult word count per hour, significantly decreased ($V = 2152$, $p = 0.037$) from the infant (*M Infant* = 3908 words, *SD Infant* = 1263.421 words) to the toddler period (*M Toddler* = 3588 words, *SD Toddler* = 1263.806

words). Similarly, the maximum turns per hour significantly decreased ($V$ = 2448, $p$ = 0.00056) from infant (*M Infant* = 125.84 turns, *SD Infant* = 45.063 turns) to toddler period (*M Toddler* = 107.1 turns, *SD Toddler* = 39.177 turns). This was also the case for the maximum child vocalisations per hour, which significantly decreased ($V$ = 2352.5, $p$ = 0.0026) from the infant (*M Infant* = 484.2 vocalisations, *SD Infant* = 172.632 vocalisations) to the toddler period (*M Toddler* = 415.1 vocalisations, *SD Toddler* = 127.881 vocalisations).

Thus, the three measures together suggest that in toddlerhood, adults are speaking less and infants also vocalise less. This reduces the amount of turn-taking children are engaged in from infancy to toddlerhood. These data differ considerably from Western samples which typically show an increase in child vocalisations and turn counts in the toddler period.

To explore the consistency of LENA measures in individuals across both time points (Q2), we ran a set of correlations. As can be seen in Figure 5.2, CTC Infant was positively related with CTC Toddler ($T$ = 4.479, $R$ = 0.447, $p$ ¡ 0.000). We found similar relationships between CVC Infant and CVC Toddler, which were positively correlated ($T$ = 2.668, $R$ = 0.285, $p$ = 0.009). This indicates that the amount of vocalisations produced by children and the amount of conversational turns that they were involved in at both time points was consistent.

We also found a significant positive relationship between AWC Infant and CTC Toddler ($T$ = 3.072, $R$ = 0.324, $p$ = 0.002). Thus, infants who were exposed to more adult words as infants engaged in more conversational turns as toddlers. In addition, CTC Infant was correlated with CVC Toddler ($T$ = 3.161, $R$ = 0.333, $p$ = 0.002), and CVC Infant was correlated with CTC Toddler ($T$ = 2.099, $R$ = 0.228, $p$ = 0.038). These relationships indicate that children who engaged in more conversational turns and produced more vocali-

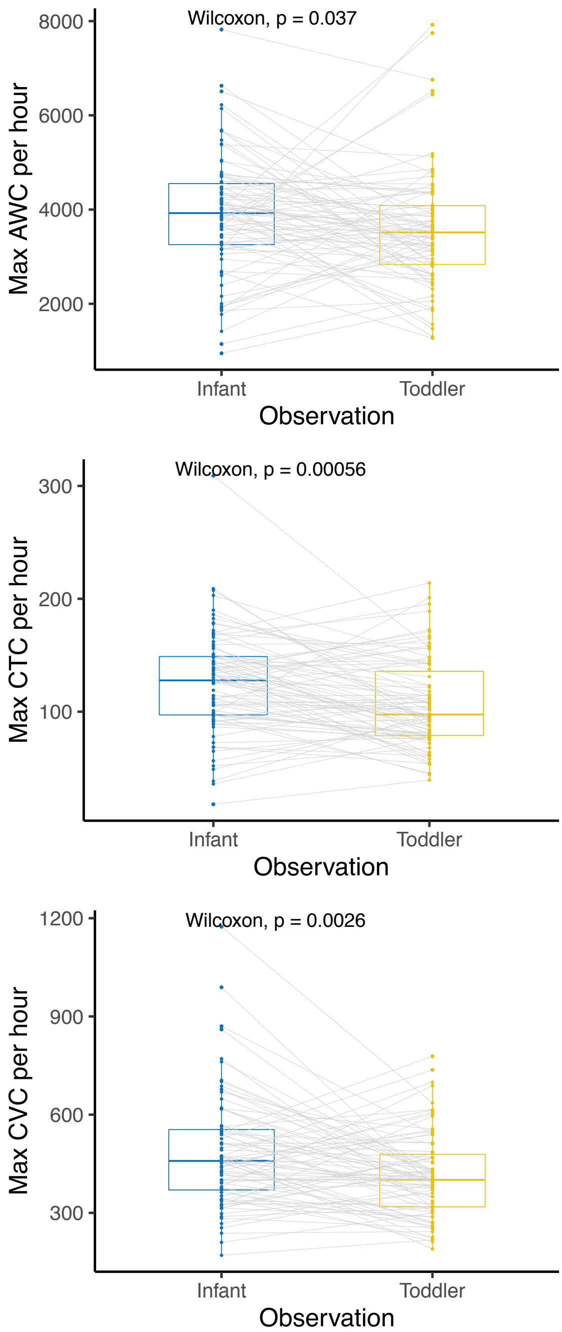


**Figure 5.1:** Max Adult word count (AWC), turn count (CTC), and child vocalization count (CVC) per hour for the infant and toddler LENA recordings. The Infant and LENA Toddler recordings are presented in blue and yellow, respectively. Individual observations across years are paired using grey lines. Results of paired Wilcoxon signed-rank tests for each LENA measure are indicated at the top of each plot.

sations as infants, also produced more vocalisations and had more conversational turns as toddlers. Finally, there was a positive correlation between AWC and CVC and CTC at both time points. This is not surprising because CTC is defined as AWC and CVC instances that happened consecutively without a pause longer than 5 seconds (see Appendix A, for a Figure A.17 containing a detailed correlation matrix with R and P values, histograms and scatterplots).

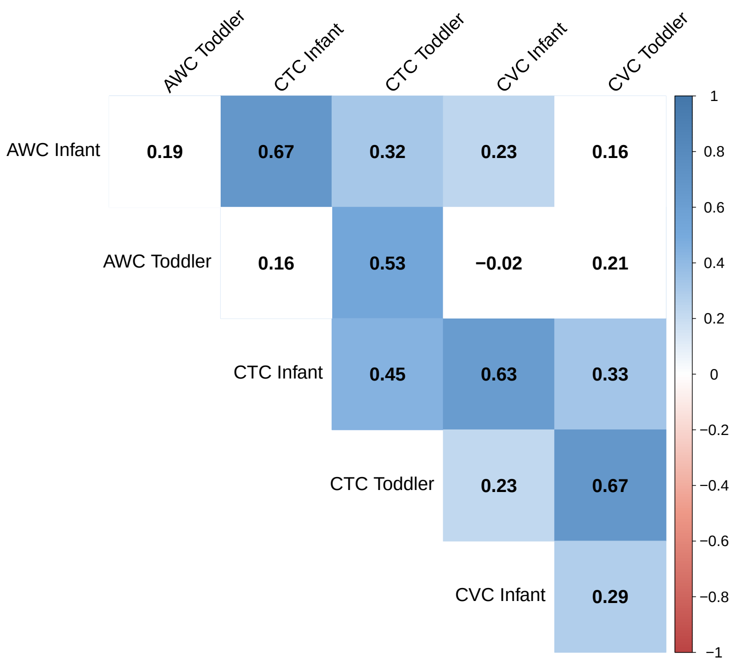


**Figure 5.2:** Uncorrected correlation matrix for LENA Infant and LENA Toddler measurements including maximum adult word count (AWC), turn count (CTC) and child vocalisations (CVC) per hour. Coloured squares display significant correlations. Positive correlations are in blue and negative correlations in red. Correlation coefficients are indicated inside each cell. Significant correlations are highlighted with colour.

The last question (Q3) was whether there were similar SES effects in our LENA sample as those reported in western samples. For instance, we wanted to examine whether higher SES children were exposed to more language input. We fit three linear models, one for each LENA measure (AWC, CTC and CVC), to the LENA data, with LENA Infant predicting LENA Toddler as a function of SES (maternal education) across LENA observations. Results showed no SES effects (see Table 5.2 for AWC F values, Table 5.3 for CTC F values and Table 5.4 for CVC F values). We also looked at relationships between SES and each LENA observation. In these linear models, SES predicted the LENA measure (in infancy or toddlerhood). We only found an SES effect on CTC in Infancy, indicating the children from higher SES families in our sample were exposed to more conversational turns than low SES children during infancy (see Figure 5.3). Note that ELP data from India showed strong SES effects (see Chapter 4), thus, we include SES in subsequent models relating ELP and LENA measures, since our outcome language measure seems to be associated with family status/ maternal education.

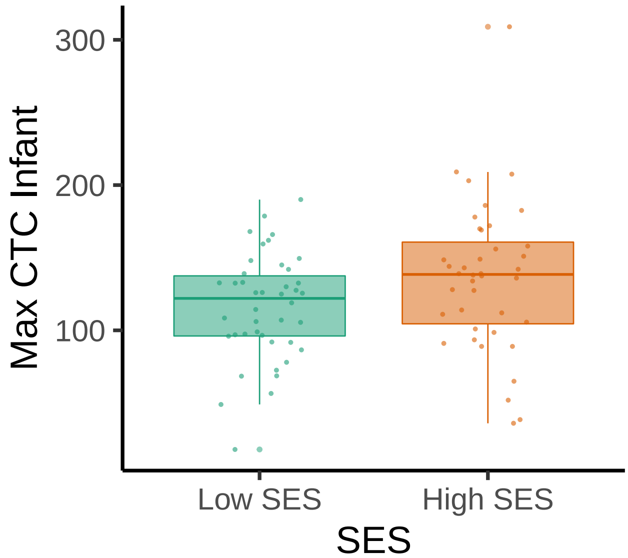


**Figure 5.3:** SES relationships for LENA maximum turn count (CTC) per hour during Infancy. Colours indicate the SES group based on maternal education.

### 5.3.2 Language Input in Relation Language Processing

To examine longitudinal relations between the LENA measures and language processing measures, we ran several regression analyses with LENA input measures, AWC and CTC, predicting ELP language processing measures at 18 months of age. All our models were fit with the LENA input measure (e.g. AWC or CTC) added as an aggregate fixed main predictor of the relevant ELP measure. We call this the "aggregated LENA". We set the LENA observations in this way because we were interested in the effect of each LENA observation (Infant or Toddler) alone, but not on the potential interactions between them which we already addressed in the previous section. This helped us simplify the models. A linear model run with the lm function of the R package (R Core Team, 2017) was used for ELP measures that contained only one or two levels such as Reinforcement, Novelty bias, Referent Selection and Retention. Variables with two levels were scaled and centered including SES, Word Type (novel versus noun) and Familiar Image Type (familiar versus highly familiar). There was no random effects structure since we only had one or two observations per participant.

For ELP measures containing several levels, such as Word Type for ELP Comprehension, we used mixed effects models, particularly glmmTMB (Brooks Mollie et al., 2017) run with the R package (R Core Team, 2017). In these models, we set Word Type and SES as fixed effects. In the random effects structure, a random intercept was nested within participant. This allowed each participant a random intercept. We set the model family to Gaussian because the proportion data was normally distributed, and thus it is expected to have a linear effect. Finally, the effect of each parameter in the models was assessed with an F test (ANOVA function from the R package), which tests whether the model terms are significant. Thus, we report F test results for our models.

#### 5.3.2.1 LENA Input and ELP Reinforcement

To measure the relationships between AWC and ELP Reinforcement, we used a linear model with the aggregated LENA AWC (AWC Infant and AWC Toddler) and SES as fixed effects predicting ELP Reinforcement. The effect of each parameter in the model was assessed with an F test that showed no significant relationships between the variables (see F Table 5.5).

To investigate the relationships between CTC and ELP Reinforcement, looking proportions to the target for ELP Reinforcement were fit with the aggregated LENA CTC (CTC Infant and CTC Toddler) and SES as predictors. The effect of each parameter in the model was assessed with an F test that revealed a significant 2-way interaction between CTC Infant and SES (see F Table 5.6). As can be seen in Figure 5.4, the number of conversational turns in infancy was negatively associated with looks to the target on ELP Reinforcement measures in Low SES children, but not in High SES children. This indicates that, when looking at highly familiar word recognition abilities, Low SES children did not benefit from conversational experiences in infancy.

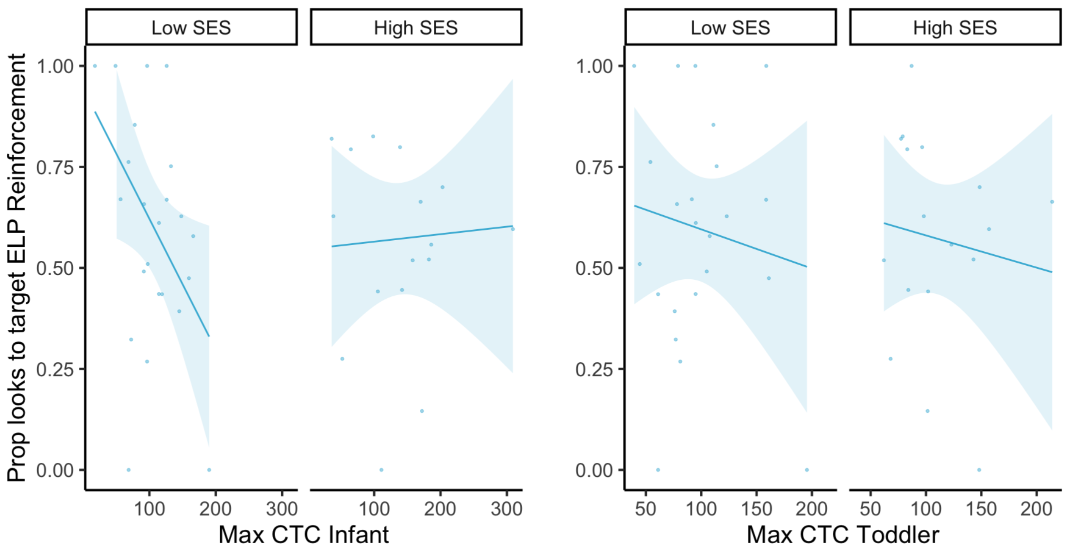


**Figure 5.4:** Relationships between CTC and ELP Reinforcement split by SES. The left panel shows CTC in infancy, the right panel shows CTC in toddlerhood.

#### 5.3.2.2 LENA Input and ELP Comprehension

To measure the relationship between AWC and ELP Comprehension, looking proportions to the target for ELP Comprehension split by word type (nouns, verbs and adjectives) and SES, were modelled using a mixed effects model. The model was fit with the aggregated LENA AWC (AWC Infant and AWC Toddler), Word Type and SES as predictors. In the random effects structure a random intercept was nested within participant, this allowed a random intercept to be fitted to each participant. The model revealed a main effect of AWC Infant and a 3-way interaction between AWC Infant, Word Type and SES (see F test values on Table 5.7). The main effect indicates a negative relationship between AWC in Infancy and looks to target on ELP Comprehension trials. However, as can be seen in Figure 5.5 this differs across words types and SES group, which is reflected on the 3-way interaction. Low SES children show negative relationships between AWC Infant and verbs and adjectives but not nouns. In contrast, High SES children show positive relationships between AWC Infant and nouns, but still negative ones between AWC Infant and verbs and adjectives. A similar pattern can be seen in relation to AWC Toddler, but the differences are not pronounced enough to reach significance.

To assess the relationships between CTC and ELP Comprehension, looking proportions to the target for ELP Comprehension split by word type (nouns, verbs and adjectives) were modelled using a mixed effects model fit with the aggregated LENA CTC (CTC Infant and CTC Toddler), Word Type and SES as predictors. In the random effects structure a random intercept was nested within participant, which allowed a random slope per participant. The effect of each parameter in the model was assessed with an F test that revealed no significant effects (see Table 5.8).

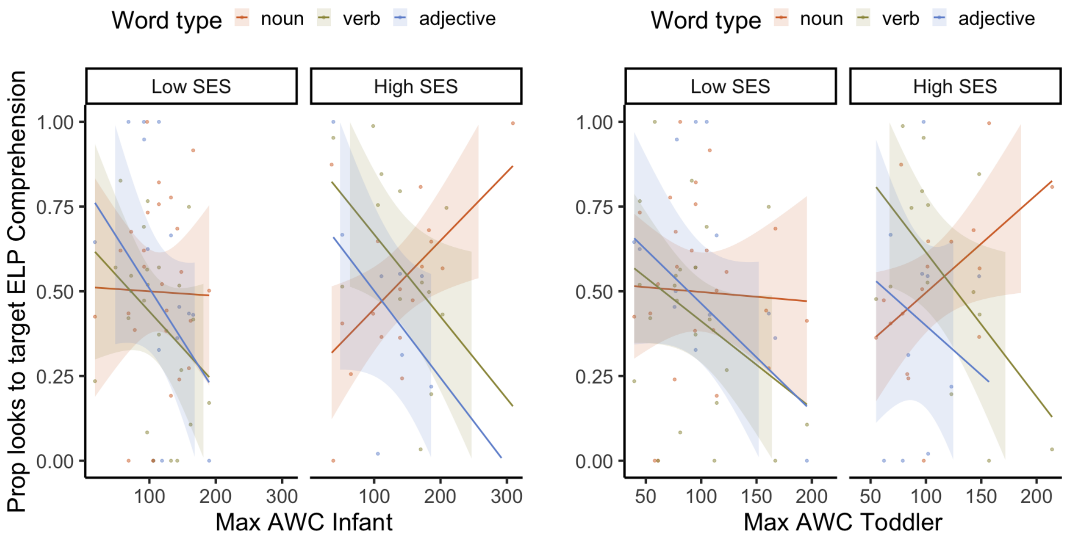


**Figure 5.5:** Relationships between AWC and ELP Comprehension split by Word Type and SES. The left panel shows AWC in infancy, the right panel shows AWC in toddlerhood.

#### 5.3.2.3 LENA Input and ELP Novelty Bias

To measure the relationships between AWC and ELP Novelty Biases, looking proportions to the novel item (i.e., ELP Novelty Bias) were fit with the aggregated LENA AWC (AWC Infant and AWC Toddler) and Familiar Image Type (familiar images that also appeared in ELP Comprehension, versus highly familiar images that also appeared in ELP Reinforcement) as predictors in a linear model. Results are reported in Table 5.9. The model did not show any significant effects.

To assess the relationships between CTC and ELP Novelty Biases, looking proportions to the novel image were fit with the aggregated LENA CTC (CTC Infancy and CTC Toddler), Familiar Image Type and SES as predictors. The effect of each parameter in the model was assessed with an F test (see Table 5.10). The model showed no significant main effects.

#### 5.3.2.4 LENA Input and ELP Referent Selection

We used a linear model to assess the relationships between AWC and ELP Referent Selection. Looking proportions to the target for ELP Referent Selection were fit with the aggregated LENA AWC (AWC Infant and AWC Toddler), Word Type (novel versus familiar noun) and SES as predictors. The effect of each parameter in the model was assessed with an F test showing a positive 2-way interaction between AWC Infant and Word Type, a 2-way interaction between AWC Toddler and Word Type and a 3-way interaction between AWC Toddler, Word Type and SES (see Table 5.11). Results are plotted in Figure 5.6. As can be seen on the left panel, more adult input in infancy led to less looks to target when this was a familiar noun, a relationship that was less pronounced when the target was novel. This was particularly the case for lower SES children, however we do not see SES effects related to AWC Infant. In contrast, the right panel shows that more adult input in toddlerhood led to less looking to novel targets. This negative relationship was stronger for higher SES children. Interestingly, high SES children also showed a positive relationship between the number of adults words they heard as toddlers and looks to familiar noun targets.

To measure relationships between CTC and ELP Referent Selection at 18-months, we ran a linear model. Looking proportions to the target for ELP Referent Selection trials were fit with the aggregated LENA CTC (CTC Infant and CTC Toddler), Word Type (novel versus familiar noun) and SES. The effect of each parameter in the model was assessed with an F test which revealed no significant effects (see Table 5.12).

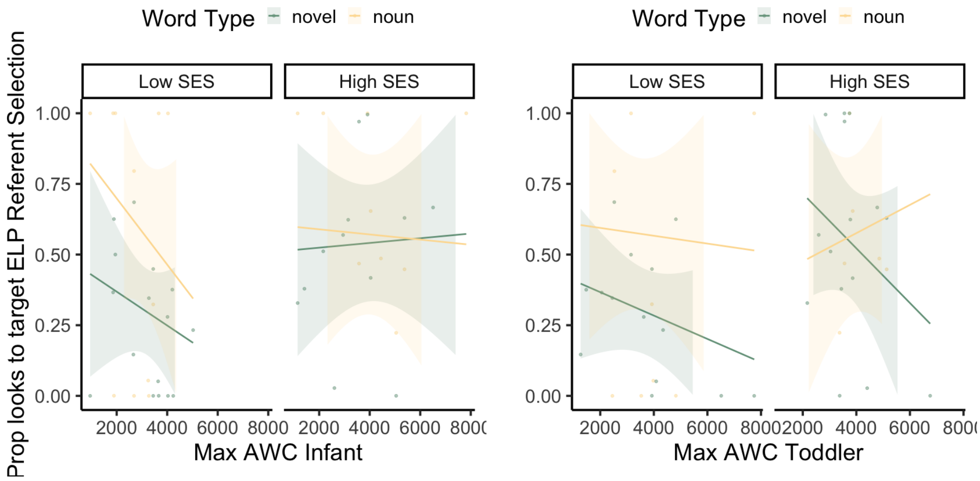


**Figure 5.6:** Relationships between AWC and ELP Referent Selection split by Word Type and SES group. The left panel shows AWC in infancy, the right panel shows AWC in toddlerhood. Different colours indicate looks to target when that was novel, versus when it was a familiar noun.

#### 5.3.2.5 LENA Input and ELP Retention

A linear model was used to measure the relationship between AWC and ELP Retention. Proportion of looks to the target on ELP Retention trials were fit with the aggregated LENA AWC (AWC Infant and AWC Toddler) and SES as predictors. The model did not show significant relationships (see Table 5.13 for the F values).

To measure possible relationships between CTC and ELP Retention, we used a linear model where looking proportions to the target were fit with the aggregated LENA CTC (CTC Infant and CTC Toddler) and SES as predictors. We assessed the effect of each parameter in the model with an F test, which revealed no significant relationships (see Table 5.14).

## 5.4 Discussion

The goal of this study was to measure the relationships between language input and the processes that support word learning in a non-Western sample, while taking SES into account. To that aim, we gathered home language input at two different time points, infancy and toddlerhood, and measures of early language processing at 18 months in a sample of Indian infants. Adult input and conversational turns were measured in children's natural environment using the LENA system. Subsequently, children's language processing abilities including speed of word processing, comprehension abilities, novelty biases, referent selection and retention of new words were measured using the Early Language Processing task.

To better understand the linguistic input that children were exposed to at home and how this changed over time, we compared the number of adult words, conversational turns that children engaged in, and vocalisations that they produced as infants and as toddlers. The first striking result is that LENA estimates in our setting were higher than those found in western samples (see Chapter 3 for LENA counts in the UK at similar ages). LENA counts were particularly high during infancy. These differences could have two interpretations. It is possible that LENA is counting more words in Awadhi due to the specific features of this language. The LENA algorithms were trained to estimate speech based on American English, thus it is possible that in Awadhi, LENA is counting syllables instead of words. An alternative is that because Indian children are surrounded by larger families, particularly during infancy when the childcare is more intense, LENA is correctly estimating the large amount of adult speech around the child. Likewise, because large families also contain a lot of speech from other children, LENA might also overestimate the child speech's as it picks up on speech from other children. The LENA system has been validated in western cultures in several languages showing high reliability, but in those contexts, speech around the child often occurs in one to one settings or in the context of small western families. We do not know how

LENA deals with large numbers of speakers, including other children producing speech around the child. We are currently investigating in more depth the particular features of the linguistic input in our sample using transcriptions of recordings (Fibla et al., In Prep). Initial analyses suggest that LENA is measuring words appropriately in the Indian context.

Our LENA results also showed a significant decrease in the number of adult words, conversational turns and vocalisations from infancy to toddlerhood. This is a surprising result – we would expect children to produce more turns and vocalisations as they become more experienced with language. Indeed, this is different from what we see in western cultures (see Chapter 3). It is possible that this decrease in children's conversational turns and vocalisations might reflect a change in childcare practices as children become older. Children's language input measures in toddlerhood were gathered when children were between 14 and 25 months of age. It is possible that at those ages, with children's increased mobility, infants spend less time surrounded by their large families and they are exposed to less adult input and they engage in fewer turns. Interestingly, even though we see a decrease in the LENA measures from infancy to toddlerhood at the group level, correlation analysis show that children exposed to more adult words during infancy engaged in more conversational turns in toddlerhood. Similarly, children who had more conversational turns in infancy vocalised more as toddlers and children who vocalised more as infants were engaged in more conversational turns as toddlers. These findings suggest language input and vocalisations influence each other to build language abilities over time.

Results exploring whether the LENA measures were related to contextual variables showed only an SES effect in conversational turns during infancy. Children from higher SES families were involved in more conversational turns than their lower SES peers during infancy. This is in line with western literature showing that higher SES parents talk more and using richer input than low SES parents (e.g., Hart & Risley, 1992, 1995; Hoff, 2003). However, the lack of SES effects in the other LENA variables is surprising, particularly given the fact the we observe strong SES effects in language processing measures gathered with the ELP task in the same sample (see Chapter 4. It is possible that SES effects are not so visible in LENA estimates because Indian families include many members. Consequently, the effects of SES based on a primary caregiver that we see in the west might be diluted given the larger quantities of speech from different speakers. This might be an explanation why we only see this effect in infancy and turn-taking, as this might be more likely to have occurred with the primary caregiver(s) at ages when child care is more intense. If this is the case, however, it is also surprising that the number of family members was not a significant predictor in our model. This could be that because Indian families in our sample are all composed of many family members and, thus, there is not much variability.

A main goal of this study was to look at the relationships between home language input gathered at two time points (infancy and toddlerhood) and children's language processing measures at 18 months accounting for SES differences. Results showed that the number of conversational turns in infancy was negatively associated with looks to the target on ELP Reinforcement measures in Low SES children, but not in High SES children. Thus, low SES children's recognition of highly familiar words did not benefit from conversational experience during infancy. It is possible that LENA estimates from turns in infancy did not contain much speech. A separate study looking at visual dynamics on dyadic interactions between Indian mothers and their infants (Forbes, Aneja, Reyes, & Spencer, 2019) has observed many instances of noise making to distract and entertain fussy infants. Although those haven not been quantified. If those indeed occur often, those interaction would lack the richness of input that has been shown in western cultures to help children's language abilities (Hoff, 2003).

Our results also showed significant negative relationships between the number of adult words the child heard as an infant and looks to target on ELP Comprehension trials. Interestingly, this differed across word types and SES. Low SES children generally did not benefit from adult input; however, high SES children benefited from adult words during infancy on noun comprehension trials. This could be an indicator that adult input in higher SES children contained more instances where children could map nouns onto objects, thereby facilitating noun comprehension at 18 months. Previous work suggests that lexical development might be highly sensitive to variation in language input (e.g., Casillas et al., 2020; Bergelson et al., 2019). In this case, it is possible that the amount of adult input but also the content of that input played a role of children's lexical abilities at 18 months. In fact, both of our results relating ELP Reinforcement or ELP Comprehension to children early language experiences, follow a similar pattern – both adult words and conversational turns show negative relationships to children's recognition abilities of nouns and highly familiar nouns. Particularly on LENA measures gathered during Infancy. This, could indicated the quantity is not enough, but that children need to engage in rich adult input and conversational experience for this to be beneficial.

Our last finding was that adult input both in infancy and toddlerhood was associated with children's disambiguation abilities on ELP Referent Selection trials. AWC toddler is positively related to referent selection on noun trials. This fits with the high SES-Comprehension effects described above. Novel word mapping relies on children's ability to map a novel word onto the correct referent in the context of a familiar object. This skill is crucial for word learning. Our results indicate that more adult input in infancy led to poorer discrimination abilities in low SES children. This negative relationship was less pronounced for higher SES children who showed a positive trend. However, larger amounts of adult input during toddlerhood led to poorer discrimination abilities for both low and high SES children. Our results show that adult input in infancy was negatively associated with children's ability to recognise familiar words. With stronger trends for low SES children. Even though we also find negative relationships between referent selection skills and adult input as toddlers in low SES children, this negative relationship is more pronounced in higher SES children. This could be explained by the positive association that we see between noun recognition abilities in the context of a novel image, and amount of adult input in toddlerhood in high SES children. It is possible that those children are experiencing a type of familiarity bias that prompts them to look more towards the familiar image. We have seen that greater amounts of adult input as toddlers in high SES leads to better recognition abilities of nouns; thus, children might be less interested in the novel item.

Overall, these results indicate that mere quantity of language input, particularly at younger ages, can hinder children's recognition abilities as well as disambiguation abilities. This results are more pronounced on the low SES children in our sample, and thus this might indicate that rather than quantity, what really matters is the content and the richness of adult input and conversational interaction. Additionally, it is possible that input from lower SES children in our sample contains more overheard speech from both adults and other children, whereas input form higher SES children could contain more instances of qualitative speech in the form of child directed speech. As in Shneidman and Goldin-Meadow (2012) children could benefit from the small amounts of rich input that they are exposed to. In contrast they could be hindered by larger amounts of overheard speech and less informative conversational experiences. These patterns show how different early language experiences can be around the globe and how the influence later abilities in different ways. We find strong SES differences in our sample regarding language processing measures, but only associated with turns in the LENA measure. This might indicate that the effect of poverty is more pronounced in the low SES group, which could translate into poorer conversational experience. However, the lack of SES effects in adult input reflects certain stability across this cultural context regarding adult speech. The SES relationships in our data regarding conversational turns, could also be an explanation why children do not benefit from language input, since poverty has been related to later developmental outcomes. In fact, early adversity in the same group of children have been reported to impact the brain networks underlying visual working memory (Wijeakumar et al., 2019), a basic skill for language development (Baddeley, 2003) and recognition abilities in the ELP task.

An impressive achievement of this project is that we managed to gather home language recordings from a large quantity of children living in a remote rural area. Moreover, a good amount of those children did the ELP task. It is remarkable that we managed to capture differences across development and SES groups using this measures in rural India.

### 5.4.1 Limitations and Future directions

A limitation of this project is that it can be hard to assess whether the performance of the younger Indian children is fully due to a delay in their language development. Some further steps could be taken to turn ELP into a valid, reliable, and robust psychometric test of individual children's language processing skills and language knowledge. For instance, some items might be especially difficult for younger children (e.g., verbs), whereas other items might be too easy for older children who show faster steeper looking time that quickly decreases, probably equating to smaller overall looking time to the target. Younger children seem to also struggle with retention trials. Something that could be done to be able to use ELP with a large age range would be to select age specific items for each task, which would match children's expected abilities and improve the psychometric robustness of ELP. To be able to do that, tests at the word level across specific age groups could be used to inform word and trial selection across ELP tasks. Moreover, an advantage of applying multiple tests in the sample (i.e., ELP and LENA) is that these could also be used to measure the psychometric properties of the ELP task, particularly its robustness as a cross-cultural tool. Children's speech data gathered with the LENA recorder, could be used as a second measure of children's language abilities, in addition to ELP.

Another limitation is that the LENA measures do not tell us how much adult speech is directed to the child. In western populations, child-directed speech has been reported to facilitate early word learning (Cartmill et al., 2013; Hoff, 2003; Rowe, 2008; Weisleder & Fernald, 2013). Data from Mayan children, who are exposed to very low amounts of CDS, also support this finding (Shneidman & Goldin-Meadow, 2012). It is possible that in large family contexts, children might be present in the scene but most speech around them would be overheard. It is the case that, because turns involve an adult addressing the child, this type of LENA measure does involve a type of child directed speech. However, we do not know what conversational turns in this setting look like and what they entail. They could be rich linguistic interactions or indicate a more or less fruitful social exchange between adult and child. We are currently looking at the characteristics of language input in rural Shivgarh by transcribing input to the child in a controlled setting during dyadic interactions in the lab, as well as in sections of the home recordings. Additional data from parent-child interactions recorded as part of this study will add more information about what characterises social exchanges between caregiver and child in early infancy in this sample.

In this study we used the maximum LENA measure per hour to predict language outcomes. This gives us an overall estimate of how much language a child is exposed to. There are other ways that these data could be explored. On one had, we could benefit from GCA (as in Donnelly & Kidd, 2021) to characterise how speech changes as a function of age. This might be a good approach for our data because we have multiple observations per participants at different ages. This would allow us to capture more fine-grained differences over development.

Another way to approach this type of data would be to characterise how often children engaged in verbal interaction with others based on time of day, household size, and number of speakers present (e.g., similarly to Casillas et al., 2020). We did explore age effects as well as overall family size. However we do not know how the speech environment of children changes across the day at different ages and how many speakers are involved. This is an interesting question because parents might engage in different activities around the child based on their age. In fact, we see that the toddler LENA measures decreased in comparison to those collected in infancy. Thus, it is possible that at older ages when children are less depended on their mother, parents engage in other activities away from the child such as farming. Thus, it is possible that at older ages, children receive less input from their primary caregivers than at younger ages. There might also be more speech from other children as has been documented in similar aged children from Mayan villages (Shneidman & Goldin-Meadow, 2012).

## 5.5 Conclusion

In conclusion, the present study provides evidence that language input, particularly in infancy, is related to later language outcomes, in particularly lexical and referent selection abilities. Importantly, we find that larger quantities of input in low SES children can hinder their language abilities at 18 months. This could be related to the specific characteristics of the language input they are exposed to.

These results set the stage for future work to measure language input in relation to language processes early in development in order to predict longer-term language outcomes, as well as working to understand how different socio-cultural environments lead to differences in language skills over development. Importantly, understanding the mechanisms that underlie these relationships could provide empirical evidence that inform intervention efforts early in development.

## 5.6 Significance Tables

**Table 5.2:** Regression results for LENA AWC and SES

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 0.03 | 1.00 | 0.03 | 0.87 | |
| AWCinfant | 2.82 | 1.00 | 2.85 | 0.10 | . |
| SES | 0.31 | 1.00 | 0.31 | 0.58 | |
| AWCinfant:SES | 0.52 | 1.00 | 0.52 | 0.47 | |
| Residuals | 77.19 | 78.00 | | | |

*Note. Main fixed effects are displayed including AWC (adult word count) Infant and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 5.3:** Regression results for LENA CTC and SES

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 0.08 | 1.00 | 0.10 | 0.75 | |
| CTCinfant | 17.55 | 1.00 | 21.70 | 0.00 | *** |
| SES | 0.18 | 1.00 | 0.22 | 0.64 | |
| CTCinfant:SES | 1.54 | 1.00 | 1.91 | 0.17 | |
| Residuals | 63.10 | 78.00 | | | |

*Note. Main fixed effects are displayed including CTC (conversational turn count) Infant and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 5.4:** Regression results for LENA CVC and SES

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 0.00 | 1.00 | 0.00 | 0.99 | |
| CVCinfant | 5.39 | 1.00 | 5.80 | 0.02 | * |
| SES | 0.28 | 1.00 | 0.30 | 0.59 | |
| CVCinfant:SES | 1.59 | 1.00 | 1.71 | 0.20 | |
| Residuals | 72.51 | 78.00 | | | |

*Note. Main fixed effects are displayed including CVC (child vocalisations count) Infant and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 5.5:** Regression results for AWC and ELP Reinforcement

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 12.32 | 1.00 | 132.23 | 0.00 | *** |
| AWCinfant | 0.03 | 1.00 | 0.34 | 0.56 | |
| AWCtoddler | 0.01 | 1.00 | 0.09 | 0.77 | |
| SES | 0.00 | 1.00 | 0.03 | 0.87 | |
| AWCinfant:SES | 0.09 | 1.00 | 0.94 | 0.34 | |
| AWCtoddler:SES | 0.02 | 1.00 | 0.19 | 0.67 | |
| Residuals | 3.45 | 37.00 | | | |

*Note. Main fixed effects are displayed including AWC (adult word count) Infant ((M = 9.76 months), AWC Toddler (M = 16.58 months) and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 5.6:** Regression results for CTC and ELP Reinforcement

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 11.78 | 1.00 | 143.87 | 0.00 | *** |
| CTCinfant | 0.25 | 1.00 | 3.01 | 0.09 | . |
| CTCtoddler | 0.00 | 1.00 | 0.03 | 0.86 | |
| SES | 0.00 | 1.00 | 0.05 | 0.82 | |
| CTCinfant:SES | 0.42 | 1.00 | 5.11 | 0.03 | * |
| CTCtoddler:SES | 0.10 | 1.00 | 1.22 | 0.28 | |
| Residuals | 3.03 | 37.00 | | | |

*Note. Main fixed effects are displayed including CTC (conversational turn count) Infant ((M = 9.76 months), CTC Toddler (M = 16.58 months) and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 5.7:** Regression results for AWC and ELP Comprehension

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 62.35 | 1.00 | 0.00 | *** |
| AWCinfant | 5.46 | 1.00 | 0.02 | * |
| AWCtoddler | 0.77 | 1.00 | 0.38 | |
| WordDType | 0.34 | 2.00 | 0.84 | |
| SES | 0.02 | 1.00 | 0.89 | |
| AWCinfant:WordDType | 1.99 | 2.00 | 0.37 | |
| AWCtoddler:WordDType | 0.43 | 2.00 | 0.81 | |
| AWCinfant:SES | 2.06 | 1.00 | 0.15 | |
| AWCtoddler:SES | 0.65 | 1.00 | 0.42 | |
| WordDType:SES | 2.40 | 2.00 | 0.30 | |
| AWCinfant:WordDType:SES | 6.13 | 2.00 | 0.05 | * |
| AWCtoddler:WordDType:SES | 1.37 | 2.00 | 0.51 | |

*Note. Main fixed effects are displayed including AWC (adult word count) Infant ((M = 9.76 months), AWC Toddler (M = 16.58 months), Word Type (nouns, verbs and adjectives) and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 5.8:** Regression results for CTC and ELP Comprehension

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 66.18 | 1.00 | 0.00 | *** |
| CTCinfant | 3.08 | 1.00 | 0.08 | . |
| CTCtoddler | 0.27 | 1.00 | 0.60 | |
| WordType | 0.79 | 2.00 | 0.68 | |
| SES | 0.00 | 1.00 | 0.99 | |
| CTCinfant:WordType | 3.67 | 2.00 | 0.16 | |
| CTCtoddler:WordType | 2.55 | 2.00 | 0.28 | |
| CTCinfant:SES | 0.07 | 1.00 | 0.80 | |
| CTCtoddler:SES | 0.22 | 1.00 | 0.64 | |
| WordType:SES | 4.22 | 2.00 | 0.12 | |
| CTCinfant:WordType:SES | 0.58 | 2.00 | 0.75 | |
| CTCtoddler:WordType:SES | 0.59 | 2.00 | 0.74 | |

*Note. Main fixed effects are displayed including CTC (conversational turn count) Infant ((M = 9.76 months), CTC Toddler (M = 16.58 months), Word Type (nouns, verbs and adjectives) and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 5.9:** Regression results for AWC and ELP Novelty Bias

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 12.98 | 1.00 | 105.34 | 0.00 | *** |
| AWCinfant | 0.03 | 1.00 | 0.24 | 0.63 | |
| AWCtoddler | 0.01 | 1.00 | 0.04 | 0.84 | |
| FamiliarImageType | 0.07 | 1.00 | 0.53 | 0.47 | |
| SES | 0.02 | 1.00 | 0.17 | 0.68 | |
| AWCinfant:FamiliarImageType | 0.07 | 1.00 | 0.56 | 0.46 | |
| AWCtoddler:FamiliarImageType | 0.02 | 1.00 | 0.15 | 0.70 | |
| AWCinfant:SES | 0.18 | 1.00 | 1.42 | 0.24 | |
| AWCtoddler:SES | 0.00 | 1.00 | 0.03 | 0.86 | |
| FamiliarImageType:SES | 0.03 | 1.00 | 0.27 | 0.60 | |
| AWCinfant:FamiliarImageType:SES | 0.29 | 1.00 | 2.36 | 0.13 | |
| AWCtoddler:FamiliarImageType:SES | 0.15 | 1.00 | 1.25 | 0.27 | |
| Residuals | 5.30 | 43.00 | | | |

*Note. Main fixed effects are displayed including AWC (adult word count) Infant ((M = 9.76 months), AWC Toddler (M = 16.58 months), Familiar Image Type (familiar images that appeared in ELP Comprehension, versus highly familiar images that appeared in ELP Reinforcement) and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 5.10:** Regression results for CTC and ELP Novelty Bias

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 13.61 | 1.00 | 106.89 | 0.00 | *** |
| CTCinfant | 0.00 | 1.00 | 0.00 | 0.95 | |
| CTCtoddler | 0.44 | 1.00 | 3.43 | 0.07 | . |
| FamiliarImageType | 0.07 | 1.00 | 0.53 | 0.47 | |
| SES | 0.03 | 1.00 | 0.26 | 0.62 | |
| CTCinfant:FamiliarImageType | 0.02 | 1.00 | 0.15 | 0.70 | |
| CTCtoddler:FamiliarImageType | 0.02 | 1.00 | 0.17 | 0.69 | |
| CTCinfant:SES | 0.00 | 1.00 | 0.01 | 0.94 | |
| CTCtoddler:SES | 0.18 | 1.00 | 1.37 | 0.25 | |
| FamiliarImageType:SES | 0.05 | 1.00 | 0.41 | 0.52 | |
| CTCinfant:FamiliarImageType:SES | 0.06 | 1.00 | 0.43 | 0.51 | |
| CTCtoddler:FamiliarImageType:SES | 0.06 | 1.00 | 0.45 | 0.51 | |
| Residuals | 5.48 | 43.00 | | | |

*Note. Main fixed effects are displayed including CTC (conversational turn count) Infant ((M = 9.76 months), CTC Toddler (M = 16.58 months), Familiar Image Type (familiar images that appeared in ELP Comprehension, versus highly familiar images that appeared in ELP Reinforcement) and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 5.11:** Regression results for AWC and ELP Referent Selection

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 11.39 | 1.00 | 78.94 | 0.00 | *** |
| AWCinfant | 0.34 | 1.00 | 2.37 | 0.13 | |
| AWCtoddler | 0.04 | 1.00 | 0.28 | 0.60 | |
| WordType | 0.30 | 1.00 | 2.09 | 0.15 | |
| SES | 0.45 | 1.00 | 3.10 | 0.09 | . |
| AWCinfant:WordType | 0.67 | 1.00 | 4.63 | 0.04 | * |
| AWCtoddler:WordType | 0.93 | 1.00 | 6.48 | 0.01 | * |
| AWCinfant:SES | 0.00 | 1.00 | 0.00 | 0.97 | |
| AWCtoddler:SES | 0.06 | 1.00 | 0.43 | 0.52 | |
| WordType:SES | 0.04 | 1.00 | 0.29 | 0.59 | |
| AWCinfant:WordType:SES | 0.26 | 1.00 | 1.82 | 0.18 | |
| AWCtoddler:WordType:SES | 0.68 | 1.00 | 4.74 | 0.04 | * |
| Residuals | 6.49 | 45.00 | | | |

*Note. Fixed effects are displayed including AWC (adult word count) Infant ((M = 9.76 months), AWC Toddler (M = 16.58 months), Word Type (familiar noun versus novel) and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 5.12:** Regression results for CTC and ELP Referent Selection

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 11.54 | 1.00 | 74.81 | 0.00 | *** |
| CTCinfant | 0.21 | 1.00 | 1.33 | 0.26 | |
| CTCtoddler | 0.00 | 1.00 | 0.00 | 0.95 | |
| TargetType | 0.31 | 1.00 | 2.01 | 0.16 | |
| SES | 0.39 | 1.00 | 2.51 | 0.12 | |
| CTCinfant:TargetType | 0.36 | 1.00 | 2.33 | 0.13 | |
| CTCtoddler:TargetType | 0.03 | 1.00 | 0.21 | 0.65 | |
| CTCinfant:SES | 0.01 | 1.00 | 0.09 | 0.77 | |
| CTCtoddler:SES | 0.03 | 1.00 | 0.20 | 0.66 | |
| TargetType:SES | 0.03 | 1.00 | 0.23 | 0.64 | |
| CTCinfant:TargetType:SES | 0.00 | 1.00 | 0.03 | 0.86 | |
| CTCtoddler:TargetType:SES | 0.01 | 1.00 | 0.05 | 0.83 | |
| Residuals | 6.94 | 45.00 | | | |

*Note. Fixed effects are displayed including CTC (conversational turn count) Infant ((M = 9.76 months), CTC Toddler (M = 16.58 months), Word Type (familiar noun versus novel) and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 5.13:** Regression results for AWC and ELP Retention

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 7.15 | 1.00 | 50.42 | 0.00 | *** |
| AWCinfant | 0.13 | 1.00 | 0.94 | 0.35 | |
| AWCtoddler | 0.21 | 1.00 | 1.51 | 0.24 | |
| SES | 0.00 | 1.00 | 0.03 | 0.86 | |
| AWCinfant:SES | 0.03 | 1.00 | 0.21 | 0.65 | |
| AWCtoddler:SES | 0.02 | 1.00 | 0.12 | 0.74 | |
| Residuals | 2.41 | 17.00 | | | |

*Note. Fixed effects are displayed including AWC (adult word count) Infant ((M = 9.76 months), AWC Toddler (M = 16.58 months) and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table 5.14:** Regression results for CTC and ELP Retention

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | sumsq | df | statistic | p.value | significance |
|---|---|---|---|---|---|
| (Intercept) | 7.52 | 1.00 | 55.92 | 0.00 | *** |
| CTCinfant | 0.10 | 1.00 | 0.76 | 0.40 | |
| CTCtoddler | 0.27 | 1.00 | 1.99 | 0.18 | |
| SES | 0.00 | 1.00 | 0.02 | 0.89 | |
| CTCinfant:SES | 0.03 | 1.00 | 0.23 | 0.64 | |
| CTCtoddler:SES | 0.00 | 1.00 | 0.01 | 0.91 | |
| Residuals | 2.29 | 17.00 | | | |

*Note. Fixed effects are displayed including CTC (conversational turn count) Infant ((M = 9.76 months), CTC Toddler (M = 16.58 months) and SES based on maternal education. Blank indicates p >.1, . indicates p <.1, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

# Chapter 6

# General Discussion

In this project we set two initial goals: 1) Develop an early language processing task that includes several measures of language processing and test it with children living in the UK and India at different times in development. 2) Measure the relationship between home language input and language processing skills at different time points in both the UK and India. These two major goals were divided into four studies which reflected smaller aims based on cultural context. First, we developed a direct measure of early language development that includes multiple measures of word learning processes. Our Early Language Processing (ELP) task was used to test language processing abilities across development in a group of children living in the UK (Chapter 2. Second, we measured the relationships between early language input and later language processes at different time points in the UK sample (Chapter 3). Third, we adapted the ELP task to a group of children from rural India in order to assess language development in an at-risk population (Chapter 4. In this sample, we also measured the effect of SES, defined by maternal education, on children's language development. Finally, we measured the relationship between early language input, SES and language processing abilities in the India sample (Chapter 5).

We effectively measured a variety of language processes early in development using a single task and related these to early language exposure. Overall, our findings indicate that we were successful; our new task shows systematic patterns of performance across multiple processes and developmental change over age. A highlight of this thesis is that the same measures were taken in two extremely different cultures, yet they show similar systematic patterns of results across cultures at both the group level and at the individual level. Thus, our work contributes a new direct measure of early language development that can be deployed in a variety of settings. Our work in India contributes to expanding the knowledge of how the environment shapes language development. Within the previous chapters our results were always discussed in the context of each culture. In the remainder of the discussion I will focus on the cross-cultural comparison.

## 6.1 Delayed early language processing in Indian children

We based the ELP task on five well established measures; speed of word processing, online comprehension, novelty biases, referent selection and retention of new words. We chose those measures because some of them are predictive of later language abilities, and because they are all essential for the word learning task. We also sought to make ELP a direct, efficient, adaptable and portable measure. For that reason we used eye-tracking techniques to capture children's looking patterns as a measure of performance on ELP. Combining multiple language processing measures in a single task allowed us to assess them in the same group of children and investigate how those processes relate to one another. The use of eye-tracking techniques allowed us to use this task on a very large range of ages, including young toddlers. Moreover, we obtained automatised fine grained measures of children's looking patterns, which allowed us to investigate children's looking dynamics on the ELP task. The fact that ELP was designed to be an automatised portable set up, allowed us to test a large sample of children in two different countries; UK and India. Moreover, ELP proved to be relatively easy to adapt to a new language and culture.

Our data shows that the ELP task accurately measured different language processes also replicating several previous finding from literature. In both UK and India, older children were faster and more accurate in speed of word processing during our ELP Reinforcement trials. Even though results from UK across ages perfectly match age effects from previous studies (see Figure 2.7 in comparison to A. Fernald et al., 1998 or Zangl & Fernald, 2007), the performance of children at the same age in India was at chance levels and it is not until older ages (above 41 months) that Indian children's looking patterns resembled those of the UK children (see Figure 4.5). Moreover, at older ages high SES Indian children showed more and faster speed of word recognition for highly familiar nouns compared to low SES same aged children. This is in line with SES differences documented in the previous literature on speed of processing in North American children from 18 and 24 months of age ((A. Fernald et al., 2013)). The delayed pattern of results seen in Indian children, as well as the differences between SES, can be seen in most ELP measures in India when compared to UK performance.

On ELP Comprehension trials, older high SES Indian children show the same pattern of word recognition across nouns, verbs and adjectives that is seen in UK children; with nouns and adjectives being easier to recognise than verbs (see Figure 2.11 for UK and Figure 4.8 for India). For nouns and adjectives we found that in both cultures older children are able to recognise adjectives almost as well as nouns, suggesting that the difference between these word types is one of difficulty. This also fits with the fact that adjectives refer to nouns. Thus, in order to learn adjectives, children need to be quite good at nouns. The advantage for nouns and adjectives over verbs is particularly interesting for two reasons. First, Awadhi is a very understudied language and, thus, frequency of appearance of different word types has never been documented, even less so in child directed speech. Word frequency has been associated with word learning (Goodman et al., 2008). One suggestion in the case of British- and American-English is that children recognise nouns earlier and faster because they appear more frequently than verbs. The similar pattern of speed of processing found in our Awadhi-learning children suggests perhaps that nouns are also more frequent in the input they receive. Another possibility is that children acquire verbs later because their referents are less concrete than nouns, making it harder for children to map their meaning onto their corresponding lexical forms (see Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005 for a discussion). If this is the case, that problem would be common across languages. However, the SES effects that we see in older Indian children points back to a role for word frequency, since low SES Indian children might be exposed to fewer word types than higher SES Indian children. It is also possible that both, UK and Indian children, had trouble recognising the 2-dimensional images of verbs used in our task. Future work analyzing the input to children in detail to have a better sense of frequency counts is needed to tease these possibilities apart.

Novelty bias was the effect that seemed most different across cultures. Although both, older children in the UK and India, showed a novelty bias, particularly in the context of highly familiar images (see Figure 2.13), some Indian children showed a familiarity bias. This implies that there was a shift in the Indian sample. Young children (e.g., between 21 and 25 months) particularly those from lower SES households, preferred to look more at the highly familiar image in the context of a novel one. However, older children showed a presence for the novel image, particularly when paired with a highly familiar one. It is interesting that in India, the degree of familiarity of the image that was paired with the novel one seemed to be highly relevant (see Figure 4.9). The shift we see from young infants to older ones could reflect several things. First, it is likely that 21 to 25 month old Indian children show a preference for the highly familiar images because they have seen it several times during the task, working as a possible priming effect. It is possible that because everything is novel to younger children, they focus on the image they have seen more often. At older ages the novelty bias we see in the context of highly familiar images could indicate lexical knowledge. Thus, when older children see an image that they know very well, they are able to shift their attention and focus on the novel one. In contrast, we see that after 29 months less well known familiar images trigger, in some cases, familiarity biases such as in the oldest low SES children. It is possible that familiarity biases in this context indicate some type of delay. In fact, if children have too much of a familiarity bias, they will not look at the novel image and thus, they will be less likely to create new word-object mappings on Referent Selection trials. This could be the opposite effect that the novelty bias literature reports in young western children (Kucker et al., 2018). However, in our case would the familiarity bias (and not the novelty bias), what prevents young children from focusing on the correct target on Referent Selection trials. The question is then, why young and low SES Indian children show a familiarity bias? A possibility might be that the set up is not fully adequate to test novelty biases in that population. Maybe everything is very novel to them and they simply look at the less novel image. Maybe repetition is the cause. We come back to this in the next paragraph.

UK children's performance on ELP Referent Selection trials replicate age effects found in previous studies (Horst & Samuelson, 2008; Bion et al., 2013) once more, we see that Indian children are delayed (see Figure 4.10 for results in India and Figure 2.14 for results in the UK). Consistently to our overall pattern of results, older high SES children are the ones to show similar responding to the UK sample. Children in India start to show disambiguation abilities between 29 and 36 months which matches the age at which we see a shift from familiarity biases to novelty biases.At even older ages, we found evidence that disambiguation in both UK and India which might be related to lexical knowledge. In India, the effect of lexical knowledge in disambiguation abilities is salient because older children show recognition of nouns on ELP Reinforcement and ELP Comprehension, which are the same familiar images we used in Reference Selection. Moreover, the novelty biases we see in older children in India, at the same age that they are able to show disambiguation abilities, might help them focus on the novel object and map the novel noun to the novel object during referent selection. The link between disambiguation skills via mutual exclusivity and lexical knowledge has been reported also in western studies (Bion et al., 2013; Kucker et al., 2018). The poorer performance of the younger Indian children on these trials may be related to their bias towards familiarity which may prevent them from focusing on the novel target during disambiguation. Since they do not know the familiar noun either (i.e., they are at chance in nouns from ELP Reinforcement and Comprehension) they are not good at recognising the familiar target and may need extra processing time focused on those objects. However, it is also possible their recognition was disrupted by the use of 2-dimensional images on a screen. A good way to investigate familiarity and novelty biases in this population would be to use the classical RSR task with 3-dimensional images (as in Horst & Samuelson, 2008). This task would be more appropriate in that sample because children are more used to see whole objects that pictures of printed 2-dimensional representations of real objects. That would help better understand the role of novelty and familiarity and its relationship to referent selection, particularly in young children.

UK children's performance on the ELP Retention trials shows some indication that 23- to 27-month-old children remembered the new word-object mappings created on referent selection trials (c.f. (Bion et al., 2013)). However this effect is not clear since retest data at 28-36 months do not show consistent retention of new words, when children theoretically should be better at it. In India, not even older high SES children are able to remember the new words. Our data from India however, diverge from previous findings as in our Indian sample even children who are older than 30-months did not show any evidence of retention. This might fit the general picture that children in India show a delay compared to Western children.

Interestingly, even though at the group level most children in UK and India did not show retention, our analysis at the individual level showed that in both cultures retention of newly formed word-object mappings was highly associated with word processing abilities and overall comprehension as measured by the Reinforcement and Comprehension trials. In India, this was the case for both low and high SES children. This indicates that children who had better language processing abilities and online word comprehension skills were also better at learning new words. Literature has shown relationships between lexical processing, vocabulary and word learning (A. Fernald et al., 2006; Lany, 2018). However, this prior work used overall vocabulary scores, rather than retention of newly formed word-object mappings, as the outcome measure. Ours is the first study to report this relationships across different processing measures and word learning. Thus is the first study to show that both speed of word processing and online comprehension are associated with remembering newly learned words across children from different cultures and low SES.

Overall then, the most striking results from our ELP measures is that the performance of older Indian children between 41 to 48 months, particularly from high SES households, reassemble that of UK children between 23 to 27 months. This is a big age gap, about 20 months. It is hard to know why younger children are not very good at ELP. A possibility is that Indian children might be inexperienced with the setup (e.g., TV or 2D images). Screens are extremely rare in Shivgarh and thus, it is possible that the lower performance seen in Indian children is due to lack of skill with the tools we used to measure language. This possibility is mitigated by, the fact that these children were part of a longitudinal study that followed their cognitive development from 6 to 48 months and included several tasks that involved stimuli presented on screens and used an eye-tracker to measure performance. Thus, children had some previous experience with the experimental setup. Alternatively, the poorer performance of the youngest children from India could reflect the impact of poverty on overall development due to nutritional deficits, economic instability, pollution, low SES, lower rates of rich parental input or fewer dyadic interactions, all of which would have consequences for language acquisition. In fact, an earlier study measuring visual working memory over development (5 to 48 months) using brain imaging techniques, conducted on a similar sample of children from Shivgarh, showed that children from low SES families (in this case calculated based on both maternal education and income) showed weaker brain activity and poorer distractor suppression in canonical working memory areas in the left frontal cortex (Wijeakumar et al., 2019).

As previously mentioned, the studies in this thesis were part of a large longitudinal study measuring the effects of poverty on child development. Convergent evidence from other tasks supports the suggestion of a cognitive delay in this population. For instance, data from a visual working memory task, indicate that Indian children (especially from low SES) have a harder time detecting change with high cognitive loads in comparison to same age UK children, and brain scans in this sample show smaller white matter brain volumes (particularly for low SES girls), than those found in same-age-children from western samples (Wijeakumar et al., In Prep). Data from the Mullen Scales of Early Learning gathered in this sample show a decrease in the receptive and expressive language scores from 6 to 9 months, with only high SES Indian girls showing an increase in expressive language with age. Overall composed scores in the Mullen scales are also higher for high SES Indian children. Relationships be-tween Mullen composed scores and physical growth gathered in this sample suggest that children with severe stunting have lower scores, and this relationship is a bit steeper for the low SES children – stunted children raised in a family with low maternal education show low Mullen scores, whereas there is little modulation of Mullen scores by stunting in families with higher maternal education. At older ages the data from Ages and Stages Questionnaire (ASQ) show that physical growth (height) is positively associated with cognition (particularly problem solving), in 18- to 21-month-old children from families with low maternal education, but there is no relationship in families with higher maternal education (J. P. Spencer, 2020).

This suggests that our language differences based on SES and age might be related to early adversity and low SES in this population. Moreover, equivalent work in non-western rural settings document the impact of early adversity in language and cognition (Milosavljevic et al., 2019; Lloyd-Fox et al., 2019). Analyses looking at ELP measures, in relation to other cognitive measures in this group of children from rural India would help to better understand this apparent language delay.

## 6.2 Cross-cultural patterns of input and input-processing relationships

In this project we also explored the role of early language input on language processing skill in UK and India. Since these two cultural contexts are extremely different, we expected to find very different results using the same measures. In particular, one might expect that as a child develops and learns more language their language production would increase and therefore the number of conversational turns they contribute to would also increase. However, a big difference between our findings in the two cultures was that turns and vocalisations from infancy to toddlerhood decreased in India (see Figure 5.1) and increased in the UK (see Figure 3.1). We also found that children in India were exposed to less adult input, particularly in infancy, compared to UK children. In general, the finding of lower rates of adult input are in line with prior findings from rural settings (Vogt & Mastin, 2013) and in indigenous communities (e.g., Casillas et al., 2020). Likewise, more input from other children would also match what has been found in non-western communities (Shneidman & Goldin-Meadow, 2012). However, this would not explain the decrease in input from infancy to toddlerhood we found in the Indian sample. Rather the cause of this finding might reside in the fact that Indian children in our sample live in large families which contain many other children. Thus it is possible that the recordings from infancy in this sample counted the speech of other children.

Another interesting finding is the same negative relationship between turns in infancy and speed of word processing, as measured on our ELP Reinforcement trials, in both UK (see Figure 3.3) and India (see Figure 5.4). However, in India this relationship was only negative for infants from low SES households. It is hard to explain why we find this negative association. It is possible that children exposed to many turns in infancy are still not ready to process that complex input. In fact, studies show that children benefit from different aspects of their input at different times of development, and young children usually benefit more from adult input (Rowe, 2012). However, there are two interesting facts related to those negative relationships in our data. First, the fact that we do not see this negative relationship between turns in infancy and ELP Reinforcement in the Indian children from high SES households indicates that turns are not "bad" for children's language development but rather that the particular features of those turns can hinder language processes. It is possible that turns in infancy are not beneficial for language because they are less focused on the linguistic aspect and more more on the social exchange or because they are produced in more noisy environments that are not optimal for early word learning. A separate study looking at visual dynamics on dyadic interactions between Indian mothers and their infants has observed many instances of noise making to distract and entertain fussy infants (Forbes et al., 2019). Although those haven not been quantified, it is possible the input to Indian infants reflect instances where caregivers try to stop children from crying, rather than provide content-rich language input. Maybe high SES Indian children are exposed to more linguistic turns in comparison to low SES Indian children. Still this does not explain why we see this negative relationship in the UK, a more homogeneous sample of middle and high SES children. In fact, in this sample we see that turns in Infancy are positively associated with ELP Comprehension at 30 months. Thus, early turns in the UK help children's online comprehension skills later on, when they are 30 months-old. Looking into the details of conversational turns in infancy and what they entail in both UK and India would help clarify what do these relationships mean.

The final cross-cultural result regarding input and language processes comes from the relationships we find in both UK and India between amount of adult input and ELP Referent Selection. In the UK the amount of adult input in infancy is positively associated with referent selection abilities (see Figure 3.5), whereas in India it is negatively associated (see Figure 3.5. This might reflect differences in adult speech across cultures. We could imagine that UK infants are exposed to more child directed speech, in comparison to Indian children who might be exposed to more overhead speech and even speech from other children. We cannot know if this is the case with our current data although. We also find a relationship between turns in infancy and referent selection in the UK that we do not see in India. This could be evidence of more child directed speech in the UK in comparison to India, since turns in the LENA system always involve the target child and an adult. Still studies show that western children are usually exposed to larger quantities of child directed speech in comparison to non-western children, particularly from some indigenous communities (Shneidman & Goldin-Meadow, 2012). We do not know if in this Indian context it is frequent to speak to children in an infant directed manner which resembles what parents do in North America or the

UK. However, we know that Indian children in our sample grow up in large families and thus, it could be expected that they are exposed to more overheard speech than UK children wo grow up in smaller families. Since lower SES families in India are usually even larger that higher SES families, this could explain why the relationship is steeper in lower SES children and still holds during toddlerhood.

## 6.3 Mechanisms and theory: The relationship between input and lexicon

There are several ways that input might influence children's language development. The relationships between language input and word learning processes include several mechanisms that facilitate learning from input, ultimately contributing to the developing lexicon. For instance, specific mechanisms might lead to faster processing or accurate responding in the different ELP trials. A possibility is that the effect of language input on children's vocabulary is based on infants' language processing efficiency because richer language experiences help children's processing skills which facilitate language growth. This would imply a discrete serial order pathway (e.g., Levelt, 1993), in which larger amounts of rich language input leads to better processing abilities, which would support better word knowledge. However, alternative pathways could include cascading or snowballing relationships (e.g., Caramazza, 1997; Apfelbaum, Blumstein, & McMurray, 2011), where different mechanisms reinforce each other such as that higher amounts of rich language input leading to faster processing which contributes to better word knowledge which would contribute, in turn, to faster word processing because well-known words are faster to process (Smolak et al., 2021).

Similarly, referent selection might be affected by the child's novelty bias as well as their knowledge, but it could also be that their knowledge of the word affects the tendency to use this knowledge as well as their novelty bias to select the referent. A study found that novel labels can disrupt visual processing for infants at an age where they are building a vocabulary, suggesting that when infants are processing labels and objects, attentional resources are shared across modalities during the developmental period of establishing a lexicon (Mather, Schafer, & Houston-Price, 2011). Simulation work using a dynamic associative model of referent selection in a two-alternative forced-choice task concluded that the processes of recognizing familiar words were not different than those that support novel words, such as fast-mapping. Moreover, processing speed was determined by experience and knowledge but also by parameters such as activation level and rate, with higher levels promoting processing speed, and competition from other known lexical items, with more lexical competition slowing speed of processing (McMurray, Horst, & Samuelson, 2012). This suggests that speed of processing derives from multiple component processes and thus, vocabulary and speed of word processing are not equivalent, because word learning does not require specialized processes and it is possible by general association learning.

Unfolding the relationships between language input and knowledge can also be complex. It is possible that specific aspects of language input contribute to better word knowledge in different ways (for a discussion see Bergelson, 2020). For example, hearing many instances of the same word across diverse contexts (input quantity), might help word knowledge because a child would be exposed to many examples of that word and thus it would be able to establish a good correspondence between the spoken word, the image that represents it and its meaning. Well established word knowledge might help children's word processing skills because the next time the child encounters that word, it will be easier to recognize. This could increase children's ability to efficiently communicate, making children more responsive to their caregivers which could contribute to increase the quality of language input that they are exposed to because caregivers might be more responsive to children that are able to communicate better with them (e.g., Choi, Nelson, Rowe, & Tager-Flusberg, 2020).

Besides the mechanisms associated supporting word learning, there are also mechanisms involved in decoding auditory input to select the target during the ELP task. In this project we tried to simplify this process by providing a lengthy familiarization phase at the start of each trial. Thus, target selection only required a decision between two alternative lexical entries that had already been activated during familiarization. Moreover, image pairs were semantically and phonologically unrelated. There is evidence of implicit naming as early as 18-months, with children directing looks towards an image based on phonological and phono-semantic priming effects (Mani & Plunkett, 2010). This suggests that phonological activation might cascade to influence the processing of visual input, and that infants might be able to extract phonological codes from the visual input.

This also raises the question whether phonological and semantic representations are activated at the same time or at different stages, after children are prompted with an image and a target audio word. A study examining the mental processes involved in lexical access in toddlers between 24- and 30-months found that language-mediated attention was influenced by both the phonological and semantic properties associated with the visual stimuli (Chow, Davies, & Plunkett, 2017). This suggests that information cascade in the lexical-semantic system is already active during early lexical development.

But what does it mean to "know" a word in the context of this task? There is evidence that infants with larger vocabularies are likely to have more refined and consolidated phonological representations which in turn impacts the activation of phonologically-related words, and hence the identification and fixation on the target (Chow et al., 2017). Thus, it is possible that older and younger children, or even two children of the same age, who make the same correct target response on a given trial are doing so on the basis of different representations. For example, some children may be able to make a correct response based on partial decoding of only the beginning of the word form while other children need the whole word to support a choice (e.g., Marslen-Wilson, 1987). However, we would expect that such differences would result in differences in processing speed. Thus, it may be that comparing children based on their complete profile of responding to easy, moderate and difficult nouns, verbs and adjectives will be most fruitful. Capturing performance patterns in computational models that can instantiate theoretical ideas about the relation between the underlying representations and these processes should also provide critical insight.

## 6.4 Limitations and future directions

In summary, we created a new direct measure of early language processing that we then adapted and used across different populations and cultures. We were able to efficiently measure language processing abilities in a large sample of children using a short 15-20 minute task. Our findings show that ELP was sensitive to developmental and individual differences. ELP captured age changes on all the five measures, with older children having a better performance on the task than younger children. Further we found systematic relationships between language input and language processing abilities in both our sample from the UK and from India. It is quite surprising that in both sites the relationships we find between language input and processes are on the same measures, ELP Reinforcement and Referent Selection and mostly involving input during infancy. This highlights the role of early language exposure on later word learning and language abilities. Moreover, this indicates the robustness of our measures which seem to be sensitive to the same variables in extremely different cultural contexts. The long-term goal of this work is to understand the processes and contextual variables that support early word learning and are predictive long-term. Doing so will require addressing some limitations of the task.

A strength of the ELP task is that combines several measures of language processes. We carefully looked at each measure as well as general relationships among them. For that we used correlation analysis which are a good first approach but do not provide many details of how the different ELP measures are related to each other either at cross-sectional timepoints or over development. Future analyses using statistical techniques that test predictive relationships between the measures could provide a much richer understanding of the relationships between the processes measured with the ELP task. Likewise, the ELP task has shown to capture several measures in one task, which should allow understanding of children's individual processing via comparisons between individual children's performance in relation to the group. That, integrated with computational modeling techniques could help understand different word learning trajectories. Our findings give some support the view that better skills in a specific language process might boost abilities in other language processes. It is possible that specific combinations of language skills lead to particular profiles of learners with different strengths. Thus, children might combine language processes in different ways when learning a word, which creates unique vocabulary pathways. Specific combinations of language processes could be beneficial while others might lead children to become late talkers. This different profiles based on particular combinations of processing skills could explain the variability we observe across children in our data, as well as in most language development research. This could be an interesting idea to test using a computational approach. Different models could represent different children varying in the strength of their processing skills. They could be tested in a task similar to ELP and we could evaluate their performance based on specific language processing combinations. This would allow to predict language outcomes and potentially identify effective interventions (see Samuelson, 2021 for a proposal). Moreover, this could be combined with longitudinal studies using the ELP measure would also contribute to our understanding long term effects of early language processes. For example, it could help identify if there are specific profiles at 18 months based on combinations of specific patterns of performance across ELP measures, that will lead to lower school achievement.

Analysis at the word level would also be useful. They would be able to show what words better capture children's abilities. That would be of particular interest in the Indian version of ELP. Analysis at the word level would help better understand how individual items in our task affected children's performance and it would give an idea of word difficulty based on group performance (i.e., if most children correctly identified a particular word).

Our input data in both UK and India could benefit from further analysis investigating what specific features in early language exposure matter for children's later processing abilities. It is possible that children in India are exposed to more overheard speech and speech from other children than children in the UK. It is also possible that the LENA turn measure is not accurate enough, particularly early in infancy. Some recent studies and reviews have questioned the validity of conversational turns measured with LENA (Ferjan Ramírez et al., 2021; Cristia et al., 2021; Cristia, Bulgarelli, & Bergelson, 2020). These questions could be addressed by transcribing our home recordings. Some research groups have developed coding schemes particularly designed to transitive naturalistic recordings (e.g., ACLEW Soderstrom et al., 2020). The downside of this approach is that it is time and resource consuming. There are some new automatised measures that capture aspects of input that might be particularly interesting in cross-cultural studies such as low child directed speech and larger amounts of overhead speech. This is the case of ALICE an open source software that extracts automatic measurement of phoneme, syllable, and word counts from child-centered daylong recordings (Räsänen et al., 2021). This would particularly fit some of the goals of this project were we worked towards more automatised measures. Automatised accurate measures of language input in combination with the efficient nature of ELP, could be a good toolkit to expand our findings longitudinally across cultures.

## 6.5 Conclusions

This thesis started with the example of Julia playing with her mother in a blossoming garden. From Julia living in the UK, we have moved to rural India, there the same example applies; Reetu is in a green garden full of colourful mangoes when a gray cat jumps from the neighbour's fence... The newly designed ELP task has shown that even though Julia and Reetu are from very different cultures they use similar tools to learn a word. Although they might apply them in different ways. Both Reetu and Julia are growing up in constant interaction with their environment which shapes their language processing abilities. The work presented here shows that even though language experiences are very different across both cultures, similar language processes might be affected by that environment. Future work will use more fine grained measures of language input to assess the impact of language experience on language development based on cultural differences in children's societies.

# References

Agostin, T. M. K., & Bain, S. K. (1997). Predicting early school success with developmental and social skills screeners. *Psychology in the Schools*, *34*(3), 219–228.

Amatuni, A., & Bergelson, E. (2017). Semantic Networks Generated from Early Linguistic Input. *Cognitive Science*.

Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). *Linking Quality and Quantity of Parental Linguistic Input to Child Language Skills: A Meta-Analysis.* Blackwell Publishing Inc.

Apfelbaum, K. S., Blumstein, S. E., & McMurray, B. (2011). Semantic priming is affected by real-time phonological competition: Evidence for continuous cascading systems. *Psychonomic Bulletin & Review*, *18*(1), 141–149.

Armstrong, R., Scott, J. G., Whitehouse, A. J., Copland, D. A., Mcmahon, K. L., & Arnott, W. (2017, 5). Late talkers and later language outcomes: Predicting the different language trajectories. *International Journal of Speech-Language Pathology*, *19*(3), 237–250.

Axelsson, E. L., & Horst, J. S. (2014). Contextual repetition facilitates word learning via fast mapping. *Acta Psychologica*, *152*, 95–99.

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*(3), 189–208.

Baqui, A. H. (2007). Newborn care in rural Uttar Pradesh. *The Indian Journal* *of Pediatrics*, *74*(3), 241–247.

Bates, E. (1993). *Comprehension and Production in Early Language Development* (Vol. 58) (No. 3-4).

Bergelson, E. (2020). The comprehension boost in early word learning: Older infants are better learners. *Child development perspectives*, *14*(3), 142–149.

Bergelson, E., & Aslin, R. N. (2017, 10). Semantic Specificity in One-Year-Olds' Word Comprehension. *Language Learning and Development*, *13*(4), 481–501. Retrieved from https://www.tandfonline.com/doi/abs/10.1080/15475441.2017.1324308

Bergelson, E., Casillas, M., Soderstrom, M., Seidl, A., Warlaumont, A. S., & Amatuni, A. (2019, 1). What Do North American Babies Hear? A large-scale cross-corpus analysis. *Developmental Science*, *22*(1).

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 of the United States of America*, *109*(9), 3253–3258.

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word- object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, *126*(1), 39–53.

Black, M. M., Walker, S. P., Fernald, L. C., Andersen, C. T., DiGirolamo, A. M., Lu, C., ... Grantham-McGregor, S. (2017). *Early childhood development coming of age: science through the life course* (Vol. 389) (No. 10064). Retrieved from http://dx.doi.org/10.1016/S0140-6736

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot. Int.*, *5*(9), 341–345.

Booth, A. E., & Waxman, S. R. (2009, 1). A horse of a different color: Specifying with precision infants' mappings of novel nouns and adjectives. *Child*

*Development*, *80*(1), 15–22.

Bowerman, M., & Choi, S. (2001). language-specific in the acquisition of spatial semantic categories. *Language acquisition and conceptual development*(3), 475.

Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, *53*, 371–399.

Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. C. (2019, 10). Consistency and Variability in Children's Word Learning Across Languages. *Open Mind*, *3*, 52–67.

Brennan, L., McDonald, J., & Shlomowitz, R. (2004). Infant feeding practices and chronic child malnutrition in the Indian states of Karnataka and Uttar Pradesh. *Economics and Human Biology*, *2*(1), 139–158.

Brent, M. R., & Siskind, J. M. (2001, 9). The role of exposure to isolated words in early vocabulary development. *Cognition*, *81*(2), B33-B44.

Brooks Mollie, E., Kristensen, K., Benthem Koen, J. V., Magnusson, A., Berg Casper, W., Nielsen, A., ... Bolker Benjamin, M. (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*, *9*(2), 378.

Brooks-Gunn, J., & Duncan, G. J. (1997). *The effects of poverty on children* (Vol. 7) (No. 2).

Brown, P., & Gaskins, S. (2014). Language acquisition and language socialization. *The Cambridge Handbook of Linguistic Anthropology*, 187–226.

Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: Infants language experience influences the development of a word learning heuristic. *Developmental Science*, *12*(5), 815–823.

Callaghan, T. C., Rochat, P., & Corbit, J. (2012). Young Children's Knowledge of the Representational Function of Pictorial Symbols: Development Across the Preschool Years in Three Cultures. *Journal of Cognition and Development*, *13*(3), 320–353.

Canault, M., Le Normand, M. T., Foudil, S., Loundon, N., & Thai-Van, H. (2016). Reliability of the Language ENvironment Analysis system (LENA™) in European French. *Behavior Research Methods*, *48*(3), 1109–1124.

Caramazza, A. (1997). How many levels of processing are there in lexical access? *Cognitive neuropsychology*, *14*(1), 177–208.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, *15*, 17–29.

Cartmill, E. A., Armstrong, 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 of the United States of America*, *110*(28), 11278–11283.

Casillas, M., Brown, P., & Levinson, S. C. (2020). Early language experience in a Papuan community. *Journal of Child Language*.

Choi, B., Nelson, C. A., Rowe, M. L., & Tager-Flusberg, H. (2020). Reciprocal influences between parent input and child language skills in dyads involving high-and low-risk infants for autism spectrum disorder. *Autism Research*, *13*(7), 1168–1183.

Chow, J., Davies, A. A., & Plunkett, K. (2017). Spoken-word recognition in 2-year-olds: The tug of war between phonological and semantic activation. *Journal of Memory and Language*, *93*, 104–134.

Connell, C. M., & Prinz, R. J. (2002). The impact of childcare and parent-child interactions on school readiness and social skills development for low-income African American children. *Journal of School Psychology*, *40*(2), 177–193.

Corriveau, K. H., Kinzler, K. D., & Harris, P. L. (2013). Accuracy trumps accent in children's endorsement of object labels. *Developmental psychology*, *49*(3), 470–479.

Cristia, A., Bulgarelli, F., & Bergelson, E. (2020). Accuracy of the language environment analysis system segmentation and metrics: A systematic review. *Journal of Speech, Language, and Hearing Research*, *63*(4), 1093–1105.

Cristia, A., Dupoux, E., Gurven, M., & Stieglitz, J. (2019). Child-Directed Speech Is Infrequent in a Forager-Farmer Population: A Time Allocation Study. *Child Development*, *90*(3), 759–773.

Cristia, A., Farabolini, G., Scaff, C., Havron, N., & Stieglitz, J. (2020, 9). Infant-directed input and literacy effects on phonological processing: Non-word repetition scores among the Tsimane'. *PLoS ONE*, *15*(9 September), e0237702.

Cristia, A., Lavechin, M., Scaff, C., Soderstrom, M., Rowland, C., Räsänen, O., ... Bergelson, E. (2021). A thorough evaluation of the Language Environment Analysis (LENA) system. *Behavior Research Methods*, *53*(2), 467–486.

Cristia, A., Seidl, A., Junge, C., Soderstrom, M., & Hagoort, P. (2014). Predicting individual variation in language from infant speech perception measures. *Child Development*, *85*(4), 1330–1345.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, and Computers*, *28*(1), 125–127.

Dale, P. S., Price, T. S., Bishop, D. V., & Plomin, R. (2003, 6). Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech, Language, and Hearing*

*Research*, *46*(3), 544–560.

Darmstadt, G. L., Kumar, V., Yadav, R., Shearer, J. C., Baqui, A. H., Awasthi, S., ... Santosham, M. (2008). Community perceptions of birth weight in rural Uttar Pradesh, India: Implications for care of low-birth-weight infants. *Journal of Perinatology*, *28*(SUPPL. 2), 53–60.

Deanda, S., Arias-Trejo, N., Poulin-Dubois, D., Zesiger, P., & Friend, M. (2016). Minimal second language exposure, SES, and early word comprehension: New evidence from a direct assessment. *Bilingualism: Language and Cognition*, *19*(01), 162–180.

Dink, J., & Ferguson, B. (2016). eyetrackingr [Computer software manual]. Retrieved from http://www.eyetracking-R.com (R package version 0.1.8)

Dollaghan, C. A., Campbell, T. F., Paradise, J. L., Feldman, H. M., Janosky, J. E., Pitcairn, D. N., & Kurs-Lasky, M. (1999). Maternal education and measures of early speech and language. *Journal of Speech, Language, and Hearing Research*, *42*(6), 1432–1443.

Donnelly, S., & Kidd, E. (2020). Individual differences in lexical processing efficiency and vocabulary in toddlers: A longitudinal investigation. *Journal of Experimental Child Psychology*, *192*, 104781. Retrieved from https://doi.org/10.1016/j.jecp.2019.104781

Donnelly, S., & Kidd, E. (2021). The Longitudinal Relationship Between Conversational Turn-Taking and Vocabulary Growth in Early Language Development. *Child Development*.

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *56*(8), 848–856.

Durham, R. E., Farkas, G., Hammer, C. S., Bruce Tomblin, J., & Catts, H. W.

(2007). Kindergarten oral language skill: A key variable in the intergenerational transmission of socioeconomic status. *Research in Social Stratification and Mobility*, *25*(4), 294–305.

Farkas, G., & Beron, K. (2004, 9). The detailed age trajectory of oral vocabulary knowledge: Differences by class and race. *Social Science Research*, *33*(3), 464–497.

Feldman, H. M., Campbell, T. F., Kurs-Lasky, M., Rockette, H. E., Dale, P. S., Colborn, D. K., & Paradise, J. L. (2005). Concurrent and predictive validity of parent reports of child language at ages 2 and 3 years. *Child Development*, *76*(4), 856–868.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., … Stiles, J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development*, *59*(5), i.

Fenson, L., Marchman, V. A., Thal, D., Dale, P. S., & Reznick, J. S. (2007). MacArthur-Bates Communicative Development Inventories: User's guide and technical manual Baltimore. *MD: Brookes*.

Ferjan Ramírez, N., Hippe, D. S., & Kuhl, P. K. (2021). Comparing Automatic and Manual Measures of Parent–Infant Conversational Turns: A Word of Caution. *Child Development*, *92*(2), 672–681.

Fernald, A., & Hurtado, N. (2006, 5). Names in frames: Infants interpret words in sentence frames faster than words in isolation. *Developmental Science*, *9*(3).

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.

Fernald, A., Marchman, V. A., & Weisleder, A. (2013). SES differences in language processing skill and vocabulary are evident at 18 months. *De-*

*velopmental Science*, *16*(2), 234–248.

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.

Fernald, A., Pinto, J. P., Swingley, D., Weinbergy, A., & McRoberts, G. W. (1998). Rapid Gains in Speed of Verbal Processing by Infants in the 2nd Year. *Psychological Science*, *9*(3), 228–231.

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When Half a Word Is Enough: Infants Can Recognize Spoken Words Using Partial Phonetic Information. *Child Development*, *72*(4), 1003–1015.

Fernald, A., Taeschner, T., Dunn, J., Papousek, M., De Boysson-Bardies, B., & Fukui, I. (1989). A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *Journal of Child Language*, *16*(3), 477–501.

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language. *Developmental psycholinguistics: On-line methods in children's language processing*, *44*, 97–135.

Fernald, L. C., Kariger, P., Hidrobo, M., & Gertler, P. J. (2012). Socioeconomic gradients in child development in very young children: Evidence from India, Indonesia, Peru, and Senegal. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(SUPPL.2), 17273–17280. Retrieved from www.pnas.org/cgi/doi/10.1073/pnas.1121241109

Fibla, L., Maniel, C., & Cristia, A. (2016). Word comprehension and multilingualism among toddlers: A study using touch screens in daycares. In *Proceedings of the joint workshop on nlp for computer assisted language learning and nlp for language acquisition* (pp. 18–23).

Fibla, L., Singh, J., Singh, V. P., Kumar, A., Samuelson, L. K., & Spencer, J. P. (In Prep). The lena system applied to the awadhi dialect.

Fisher, E. L. (2017, 10). *A systematic review and meta-analysis of predictors of expressive-language outcomes among late talkers* (Vol. 60) (No. 10). American Speech-Language-Hearing Association.

Forbes, S., Aneja, P., Reyes, L. D., & Spencer, J. (2019). Understandings caregivers' visual experience and infants' cognitive outcomes: A cross-cultural study. In *Xvi european congress of psychology* (pp. 569–569).

Ford, M., Baer, C. T., Xu, D., Yapanel, U., & Gray, S. (2008). The LENA™ language environment analysis system.

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, *44*(3), 677–694.

Friend, M., & Keplinger, M. (2003). An infant-based assessment of early lexicon acquisition. *Behavior Research Methods, Instruments, and Computers*, *35*(2), 302–309.

Friend, M., & Keplinger, M. (2008). Reliability and validity of the Computerized Comprehension Task (CCT): data from American English and Mexican Spanish infants*. *Journal of Child Language*, *35*(1), 77–98.

Friend, M., Smolak, E., Liu, Y., Poulin-Dubois, D., & Zesiger, P. (2018, 7). A cross-language study of decontextualized vocabulary comprehension in toddlerhood and kindergarten readiness. *Developmental Psychology*, *54*(7), 1317–1333.

Friend, M., Smolak, E., Patrucco-Nanchen, T., Poulin-Dubois, D., & Zesiger, P. (2019, 1). Language status at age 3: Group and individual prediction from vocabulary comprehension in the second year. *Developmental Psychology*, *55*(1), 9–22.

Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Oller, D. K., ... 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.

Gilkerson, J., Richards, J. A., Warren, S. F., Oller, D. K., Russo, R., & Vohr, B. (2018). Language experience in the second year of life and language outcomes in late childhood. *Pediatrics*, *142*(4).

Gilkerson, J., Zhang, Y., Xu, D., Richards, J. A., Xu, X., Jiang, F., ... Topping, K. (2015, 4). Evaluating language environment analysis system performance for Chinese: A pilot study in Shanghai. *Journal of Speech, Language, and Hearing Research*, *58*(2), 445–452.

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language learning and development*, *1*(71), 23–64.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*(1), 23–45.

Golinkoff, R. M., Hoff, E., Rowe, M. L., Tamis-LeMonda, C. S., & Hirsh-Pasek, K. (2019, 5). *Language Matters: Denying the Existence of the 30-Million-Word Gap Has Serious Consequences* (Vol. 90) (No. 3). Blackwell Publishing Inc.

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.

Greenwood, C. R., Bourque, K., & Buzhardt, J. (2011, 2). Assessing children's home language environments using automatic speech recognition technology. *journals.sagepub.com*, *32*(2), 83–92. Retrieved from https://www.researchgate.net/publication/254081850

Hackman, D. A., & Farah, M. J. (2009). *Socioeconomic status and the developing brain* (Vol. 13) (No. 2).

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.

Hart, B., & Risley, T. R. (1992). American Parenting of Language-Learning Children: Persisting Differences in Family-Child Interactions Observed in Natural Home Environments. *Developmental Psychology*, *28*(6), 1096–1105.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Paul H Brookes Publishing.

Hartig, F. (2021). Dharma: Residual diagnostics for hierarchical (multi-level / mixed) regression models [Computer software manual]. Retrieved from `http://florianhartig.github.io/DHARMa/` (R package version 0.4.1)

Heath, S. B. (1982). What no bedtime story means: Narrative skills at home and school. *Language in society*, 49–76.

Heath, S. B. (1983). *Ways with Words: Language, Life and Work in Communities and Classrooms*. Cambridge: Cambridge University Press.

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009, 6). Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition? *Psychological Science*, *20*(6), 729–739.

Hirsh-Pasek, K., Adamson, L. B., Bakeman, R., Owen, M. T., Golinkoff, R. M., Pace, A., … Suma, K. (2015). The Contribution of Early Communication Quality to Low-Income Children's Language Success. *Psychological Science*, *26*(7), 1071–1083.

Hoff, E. (2003). The Specificity of Environmental Influence: Socioeconomic Status Affects Early Vocabulary Development Via Maternal Speech.

*Child Development*, *74*(5), 1368–1378.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review*, *26*(1), 55–88.

Hoff, E., Laursen, B., & Tardif, T. (2002). Socioeconomic status and parenting. In *Mh bornstein (ed) handbook of parenting vol. ii* (pp. 161–188). Mahwah, New Jersey: Lawrence Erlbaum.

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, *73*(2), 418–433.

Hoff, E., & Tian, C. (2005). Socioeconomic status and cultural influences on language. *Journal of Communication Disorders*, *38*(4 SPEC. ISS.), 271–278.

Hoff-Ginsberg, E. (1991). Mother-Child Conversation in Different Social Classes and Communicative Settings. *Child Development*, *62*(4), 782–796.

Horst, J. S., & Hout, M. C. (2016, 12). The Novel Object and Unusual Name (NOUN) Database: A collection of novel images for use in experimental research. *Behavior Research Methods*, *48*(4), 1393–1409. Retrieved from https://link.springer.com/article/10.3758/s13428-015-0647-3

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, *13*(2), 128–157.

Horst, J. S., Samuelson, L. K., Kucker, S. C., & McMurray, B. (2011, 2). What's new? Children prefer novelty in referent selection. *Cognition*, *118*(2), 234–244.

Houwer, A. D. (2011). Language input environments and language development in bilingual acquisition. *Applied Linguistics Review*, *2*(2011), 221–240.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008a). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, *11*(6), F31-F39.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008b). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science*, *11*(6), F31-F39.

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.

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive Psychology*, *45*(3), 337–374.

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.

Jones, G., & Rowland, C. F. (2017, 11). 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.

Jusczyk, P. W. (1999, 9). *How infants begin to extract words from speech* (Vol. 3) (No. 9). Trends Cogn Sci.

Kalashnikova, M., Mattock, K., & Monaghan, P. (2015). The effects of linguistic experience on the flexible use of mutual exclusivity in word learning. *Bilingualism*, *18*(4), 626–638.

Kapalková, S., Polišenská, K., & Vicenová, Z. (2013, 1). Non-word repetition performance in Slovak-speaking children with and without SLI: Novel scoring methods. *International Journal of Language and Communication Disorders*, *48*(1), 78–89.

Kartushina, N., & Mayor, J. (2019). Word knowledge in six- To nine-month-old Norwegian infants? Not without additional frequency cues. *Royal Society Open Science*, *6*(9).

Kelly, Y., Sacker, A., Del Bono, E., Francesconi, M., & Marmot, M. (2011). What role for the home learning environment and parenting in reducing the socioeconomic gradient in child development? Findings from the Millennium Cohort Study. *Archives of Disease in Childhood*, *96*(9), 832–837.

*Key Indicators of Household Social Consumption on Education in India. NSS 75th Round* (Tech. Rep.). (2018). New Delhi, India: Government of India Ministry of Statistics and Programme Implementation.

Kucker, S. C., McMurray, B., & Samuelson, L. K. (2018). Too Much of a Good Thing: How Novelty Biases and Vocabulary Influence Known and Novel Referent Selection in 18-Month-Old Children and Associative Learning Models. *Cognitive Science*, *42*, 463–493.

Kucker, S. C., McMurray, B., & Samuelson, L. K. (2020). Sometimes it is better to know less: How known words influence referent selection and retention in 18- to 24-month-old children. *Journal of Experimental Child Psychology*, *189*.

Kucker, S. C., & Samuelson, L. K. (2012, 5). The First Slow Step: Differential Effects of Object and Word-Form Familiarization on Retention of Fast-Mapped Words. *Infancy*, *17*(3), 295–323.

Kumar, V., Mohanty, S., Kumar, A., Misra, R. P., Santosham, M., Awasthi, S., ... Darmstadt, G. L. (2008). Effect of community-based behaviour change management on neonatal mortality in Shivgarh, Uttar Pradesh, India:

a cluster-randomised controlled trial. *The Lancet*, *372*(9644), 1151–1162.

*Language. India, States and Union Territories* (Vol. 46; Tech. Rep. No. 4). (2011). New Delhi, India: Government of India.

Lany, J. (2018, 5). Lexical-processing efficiency leverages novel word learning in infants and toddlers. *Developmental Science*, *21*(3), e12569.

Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics*, *32*(1), 69–92.

Lee, V. E., & Burkam, D. T. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. ERIC.

Legacy, J., Zesiger, P., Friend, M., & Poulin-Dubois, D. (2018). Vocabulary size and speed of word recognition in very young French-English bilinguals: A longitudinal study. *Bilingualism (Cambridge, England)*, *21*(1), 137–149.

Leroy, J. L., Gadsden, P., & Guijarro, M. (2012). The impact of daycare programmes on child health, nutrition and development in developing countries: A systematic review. *Journal of Development Effectiveness*, *4*(3), 472–496.

Levelt, W. J. (1993). *Speaking: From intention to articulation* (Vol. 1). MIT press.

Lewis, M., Cristiano, V., Lake, B. M., Kwan, T., & Frank, M. C. (2020). The role of developmental change and linguistic experience in the mutual exclusivity effect. *Cognition*, *198*.

Lieven, E. (1994). Crosslinguistic and crosscultural aspects of language addressed to children. In C. Gallaway & B. Richards (Eds.), *Input and interaction in language acquisition* (pp. 56–73). Cambridge: Cambridge University Press.

Lieven, E., Salomo, D., & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: A usage-based analysis.

Lieven, E., & Stoll, S. (2013). Early communicative development in two cultures: A comparison of the communicative environments of children from two cultures. *Human Development*, *56*(3), 178–206.

Lloyd-Fox, S., Blasi, A., McCann, S., Rozhko, M., Katus, L., Mason, L., ... Elwell, C. E. (2019, 9). Habituation and novelty detection fNIRS brain responses in 5- and 8-month-old infants: The Gambia and UK. *Developmental Science*, *22*(5), e12817.

Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P., & Makowski, D. (2021). Assessment, testing and comparison of statistical models using r. *Journal of Open Source Software*, *6*(59), 3112.

MacWhinney, B. (2000). *The Childes Project: The Database*. Psychology Press.

Mahr, T., & Edwards, J. (2018, 11). Using language input and lexical processing to predict vocabulary size. *Developmental Science*, *21*(6).

Mani, N., & Plunkett, K. (2010). In the infant's mind's ear: Evidence for implicit naming in 18-month-olds. *Psychological science*, *21*(7), 908–913.

Maniel, C. (2016). *Effet du statut socio-économique sur la compréhension de mots des jeunes enfants à Paris* (Unpublished master's thesis). Université Pierre et Marie Curie.

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.

Marchman, V. A., Ashland, M. D., Loi, E. C., Adams, K. A., Fernald, A., & Feldman, H. M. (2019). Predictors of early vocabulary growth in children born preterm and full term: A study of processing speed and medical complications. *Child Neuropsychology*, *25*(7), 943–963.

Marchman, V. A., & Fernald, A. (2008). Speed of word recognition and vocabulary knowledge in infancy predict cognitive and language outcomes in later childhood: Fast-Track Report. *Developmental Science*, *11*(3), 9–16.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121–157.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word-recognition. *Cognition*, *25*(1-2), 71–102.

Mather, E. (2013). Novelty, attention, and challenges for developmental psychology. *Frontiers in Psychology*, *4*(AUG).

Mather, E., Schafer, G., & Houston-Price, C. (2011). The impact of novel labels on visual processing during infancy. *British Journal of Developmental Psychology*, *29*(4), 783–805.

McDonald, M., Kwon, T., Kim, H., Lee, Y., & Ko, E. S. (2021, 3). Evaluating the language environment analysis system for Korean. *Journal of Speech, Language, and Hearing Research*, *64*(3), 792–808.

McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological review*, *119*(4), 831.

Meints, K., Plunkett, K., & Harris, P. L. (1999). When does an ostrich become a bird? The role of typicality in early word comprehension. *Developmental psychology*, *35*(4), 1072–1078.

Merriman, W. E., & Schuster, J. M. (1991, 12). Young Children's Disambiguation of Object Name Reference. *Child Development*, *62*(6), 1288–1301.

Milosavljevic, B., Vellekoop, P., Maris, H., Halliday, D., Drammeh, S., Sanyang, L., … Lloyd-Fox, S. (2019, 9). Adaptation of the Mullen Scales of Early

Learning for use among infants aged 5- to 24-months in rural Gambia. *Developmental Science*, *22*(5), e12808.

Mirman, D. (2014). *Growth Curve Analysis and Visualization Using R*. New York: Chapman and Hall/CRC.

Mullen, E. M., & Others. (1995). *Mullen scales of early learning*. AGS Circle Pines, MN.

Newman, R. S., Rowe, M. L., & Bernstein Ratner, N. (2016). Input and uptake at 7 months predicts toddler vocabulary: The role of child-directed speech and infant processing skills in language development. *Journal of Child Language*, *43*(5), 1158–1173.

Noble, K. G., Grieve, S. M., Korgaonkar, M. S., Engelhardt, L. E., Griffith, E. Y., Williams, L. M., & Brickman, A. M. (2012, 10). Hippocampal volume varies with educational attainment across the life-span. *Frontiers in Human Neuroscience*(OCTOBER 2012), 1–26.

Ochs, E., Schieffelin, B., et al. (1984). Language acquisition and socialization. *Culture theory: Essays on mind, self, and emotion*, 276–320.

Orena, A. J., Byers-Heinlein, K., & Polka, L. (2019). Reliability of the Language Environment Analysis (LENA) in French-English Bilingual Speech. *eScholarship@McGill: Discover theses, dissertations, articles, and more*, *1*(514).

Pakulak, E., & Neville, H. J. (2010, 12). *Proficiency differences in syntactic processing of monolingual native speakers indexed by event-related potentials* (Vol. 22) (No. 12).

Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal Correlates of Growth in Toddler Vocabulary Production in Low-Income Families. *Child Development*, *76*(4), 763–782.

Peute, B., & Casillas, M. (2021). Early consonant production in Tseltal and

Yélî Dnye.

Piazzalunga, S., Previtali, L., Pozzoli, R., Scarponi, L., & Schindler, A. (2019, 5). An articulatory-based disyllabic and trisyllabic Non-Word Repetition test: reliability and validity in Italian 3- to 7-year-old children. *Clinical Linguistics and Phonetics*, *33*(5), 437–456.

Pye, C. (1992). Language loss among the Chilcotin. *International Journal of the Sociology of Language*, *93*(1), 75–86.

Quigley, J., & McNally, S. (2013). Mother-infant interaction corpus: Infant siblings at risk of autism spectrum disorders and typically developing infants.

Quine, W. V. O. (1960). *Word and object.* Cambridge MA: MIT Press.

R Core Team. (2017). R: A language and environment for statistical computing. *R Found. Stat. Comput. Vienna, Austria..* Retrieved from `www.R-project.org`

Ramey, C. T., & Ramey, S. L. (2004). *Early learning and school readiness: Can early intervention make a difference?* (Vol. 50) (No. 4). Wayne State University Press.

Räsänen, O., Seshadri, S., Lavechin, M., Cristia, A., & Casillas, M. (2021, 4). ALICE: An open-source tool for automatic measurement of phoneme, syllable, and word counts from child-centered daylong recordings. *Behavior Research Methods*, *53*(2), 818–835.

Reese, E., & Read, S. (2000). Predictive validity of the New Zealand MacArthur Communicative Development Inventory: Words and Sentences. *Journal of Child Language*, *27*(2), 255–266.

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.

Rescorla, L. (2011, 11). Late talkers: Do good predictors of outcome exist? *Developmental Disabilities Research Reviews*, *17*(2), 141–150.

Ring, E. D., & Fenson, L. (2000). The correspondence between parent report and child performance for receptive and expressive vocabulary beyond infancy. *First Language*, *20*(59), 141–159.

Roberts, J. E., Burchinal, M., & Durham, M. (1999). Parents' report of vocabulary and grammatical development of African American preschoolers: Child and environmental associations. *Child Development*, *70*(1), 92–106.

Rodriguez, E. T., & Tamis-Lemonda, C. S. (2011, 7). Trajectories of the home learning environment across the first 5years: Associations with children's vocabulary and literacy skills at prekindergarten. *Child Development*, *82*(4), 1058–1075.

Romeo, R. R., Segaran, J., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., … Gabrieli, J. D. (2018). Language exposure relates to structural neural connectivity in childhood. *Journal of Neuroscience*, *38*(36), 7870–7877.

Rosemberg, C. R., & Alam, F. (2021, 1). Socioeconomic disparities in the comprehension of lexical categories. A study with Spanish-speaking Argentinian toddlers. *European Journal of Psychology of Education*, 1–20.

Rost, G. C., & McMurray, B. (2009, 3). Speaker variability augments phonological processing in early word learning. *Developmental Science*, *12*(2), 339–349.

Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*, *35*(1), 185–205.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech vocabulary development. *Child Development*, *83*(5), 1762–1774.

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.

Rowe, M. L., & Snow, C. E. (2020, 1). Analyzing input quality along three dimensions: Interactive, linguistic, and conceptual. *Journal of Child Language*, *47*(1), 5–21.

Rowe, M. L., & Weisleder, A. (2020). Language Development in Context. *Annual Review of Developmental Psychology*, *2*, 201–223.

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 of the United States of America*, *112*(41), 12663–12668.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926–1928.

Sahithya, B. R., Manohari, S. M., & Vijaya, R. (2019). Parenting styles and its impact on children–a cross cultural review with a focus on India. *Mental Health, Religion and Culture*, *22*(4), 357–383.

Saksena, B. (1971). *Evolution of Awadhi (a Branch of Hindi)* (2nd Editio ed.). Delhi: Motilal Banarsidass Publ.

Samuelson, L. K. (2021). Toward a Precision Science of Word Learning: Understanding Individual Vocabulary Pathways. *Child Development Perspectives*, *15*(2), 117–124.

Samuelson, L. K., Kucker, S. C., & Spencer, J. P. (2017). Moving Word Learning to a Novel Space: A Dynamic Systems View of Referent Selection and Retention. *Cognitive Science*, *41*, 52–72.

Sarsour, K., Sheridan, M., Jutte, D., Nuru-Jeter, A., Hinshaw, S., & Boyce, T. W.

(2011). Family socioeconomic status and child executive functions: The roles of language, home environment, and single parenthood. *Journal of the International Neuropsychological Society*, *17*(1), 120–132.

Scaff, C., Fibla, L., & Cristia, A. (in press). Factors structuring lexical development in toddlers: The effects of parental education, language exposure, and age. *Journal of Child Language*.

Scarborough, H. S. (2009). Connecting early language and literacy to later reading (dis) abilities: Evidence, theory, and practice. In *Approaching difficulties in literacy development: Assessment, pedagogy and programmes* (pp. 23–28).

Schwarz, I.-c., Botros, N., Lord, A., Marcusson, A., Tidelius, H., & Marklund, E. (2017). The LENA TM system applied to Swedish : Reliability of the Adult Word Count estimate CLINTEC , Speech Pathology Unit , Karolinska Institutet , Stockholm , Sweden. *Interspeech*, 2088–2092.

Shneidman, L. A., & Goldin-Meadow, S. (2012, 9). Language input and acquisition in a Mayan village: How important is directed speech? *Developmental Science*, *15*(5), 659–673.

Smith-Hefner, N. (1988). The Linguistic Socialization of Javanese Children in Two Communities. *Anthropological linguistics*, *30*(2), 166–198.

Smolak, E., Hendrickson, K., Zesiger, P., Poulin-Dubois, D., & Friend, M. (2021). Visual and haptic responses as measures of word comprehension and speed of processing in toddlers: Relative predictive utility. *Journal of Experimental Child Psychology*, *203*, 105032.

Snow, C. E. (1972, 6). Mothers' Speech to Children Learning Language. *Child Development*, *43*(2), 549.

Snow, C. E. (1977). The development of conversation between mothers and babies. *Journal of child language*, *4*(1), 1–22.

Soderstrom, M., Casillas, M., Bergelson, E., Rosemberg, C. R., Alam, F., Warlaumont, A. S., & Bunce, J. P. (2020). Developing A Cross-Cultural Annotation System and MetaCorpus for Studying Infants' Real World Language Experience. *PsyArXiv Preprints*, *September*.

Song, L., Spier, E. T., & Tamis-Lemonda, C. S. (2014). Reciprocal influences between maternal language and children's language and cognitive development in low-income families. *Journal of Child Language*, *41*(2), 305–326.

Spencer, E. J., & Schuele, C. M. (2012). An examination of fast mapping skills in preschool children from families with low socioeconomic status. *Clinical Linguistics and Phonetics*, *26*(10), 845–862.

Spencer, J. P. (2020, June). personal communication.

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.

Spiegel, C., & Halberda, J. (2011). Rapid fast-mapping abilities in 2-year-olds. *Journal of Experimental Child Psychology*, *109*(1), 132–140.

Squires, J., Bricker, D., Clifford, J., Murphy, K., Hoselton, R., Potter, L., ... Farrell, J. (2009). *Ages & Stages Questionnaires ® A Parent-Completed Child Monitoring System THIRD EDITION ASQ-3™ CD-ROM* (Tech. Rep.).

Swingley, D., & Humphrey, C. (2018, 7). Quantitative Linguistic Predictors of Infants' Learning of Specific English Words. *Child Development*, *89*(4), 1247–1267.

Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the "noun bias" in context: A comparison of English and Mandarin. *Child Development*, *70*(3), 620–635.

The GIMP Development Team. (n.d.). *Gimp*. Retrieved from https://www

.gimp.org

Tomasello, M. (2019). *Becoming human: A theory of ontogeny.* Cambridge MA: Belknap Press.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child development*, *57*(6), 1454–1463.

Tomasello, M., & Mervis, C. B. (1994). *The instrument is great, but measuring comprehension is still a problem* (Vol. 59) (No. 5). Wiley Online Library.

Twomey, K. E., Ranson, S. L., & Horst, J. S. (2014). That's More Like It: Multiple Exemplars Facilitate Word Learning. *Infant and Child Development*, *23*(2), 105–122.

Vernon-Feagans, L., Bratsch-Hines, M., Reynolds, E., & Willoughby, M. (2020). How Early Maternal Language Input Varies by Race and Education and Predicts Later Child Language. *Child Development*, *91*(4), 1098–1115.

Vogt, P., & Mastin, J. D. (2013). Rural and urban differences in language socialization and early vocabulary development in Mozambique. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *35*(35), 882–887.

Wade, M., Browne, D. T., Madigan, S., Plamondon, A., & Jenkins, J. M. (2014). Normal birth weight variation and children's neuropsychological functioning: Links between language, executive functioning, and theory of mind. *Journal of the International Neuropsychological Society*, *20*(9), 909–919.

Wagenmakers, E. J., & Farrell, S. (2004). AIC model selection using Akaike weights. *Psychonomic Bulletin and Review*, *11*(1), 192–196.

Wang, Y., Williams, R., Dilley, L., Houston, D. M., Surgery, N., & Lansing, E. (2020). A meta-analysis of the predictability of LENA™ automated measures for child language development. *Developmental Review*, *57*, 100921.

Waxman, S., Fu, X., Arunachalam, S., Leddon, E., Geraghty, K., & Song, H. j. (2013, 9). Are nouns learned before verbs? Infants provide insight into a long-standing debate. *Child Development Perspectives*, *7*(3), 155–159.

Weber, A. M., Marchman, V. A., DIop, Y., & Fernald, A. (2018). Validity of caregiver-report measures of language skill for Wolof-learning infants and toddlers living in rural African villages. *Journal of Child Language*, *45*(4), 939–958.

Weisleder, A., & Fernald, A. (2013). Talking to Children Matters: Early Language Experience Strengthens Processing and Builds Vocabulary. *Psychological Science*, *24*(11), 2143–2152.

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.

Wells, G., & Bridges, A. (1981). *Learning through interaction: volume 1: the study of language development* (Vol. 1). Cambridge University Press.

Werker, J. F., & Curtin, S. (2005, 4). PRIMIR: A Developmental Framework of Infant Speech Processing. *Language Learning and Development*, *1*(2), 197–234.

Wijeakumar, S., Forbes, S. H., Magnotta, V. A., Singh, V. P., Kumar, A., & Spencer, J. P. (In Prep). Stunting in the first year of life is associated with changes in the dorsal and ventral attention networks.

Wijeakumar, S., Kumar, A., Delgado Reyes, L. M., Tiwari, M., & Spencer, J. P. (2019, 9). Early adversity in rural India impacts the brain networks underlying visual working memory. *Developmental Science*, *22*(5), e12822.

Yurovsky, D., Bion, R. A., Smith, L. B., & Fernald, A. (2012). Mutual Exclusivity and Vocabulary Structure. *Proceedings of the 34th annual meeting of the Cognitive Science Society*, 1197–1202.

Zangl, R., & Fernald, A. (2007, 6). Increasing Flexibility in Children's Online Processing of Grammatical and Nonce Determiners in Fluent Speech. *Language Learning and Development*, *3*(3), 199–231.

Zimmerman, F. J., Gilkerson, J., Richards, J. A., Christakis, D. A., Xu, D., Gray, S., & Yapanel, U. (2009). Teaching by listening: The importance of adult-child conversations to language development. *Pediatrics*, *124*(1), 342–349.

# Appendix A

# Appendix

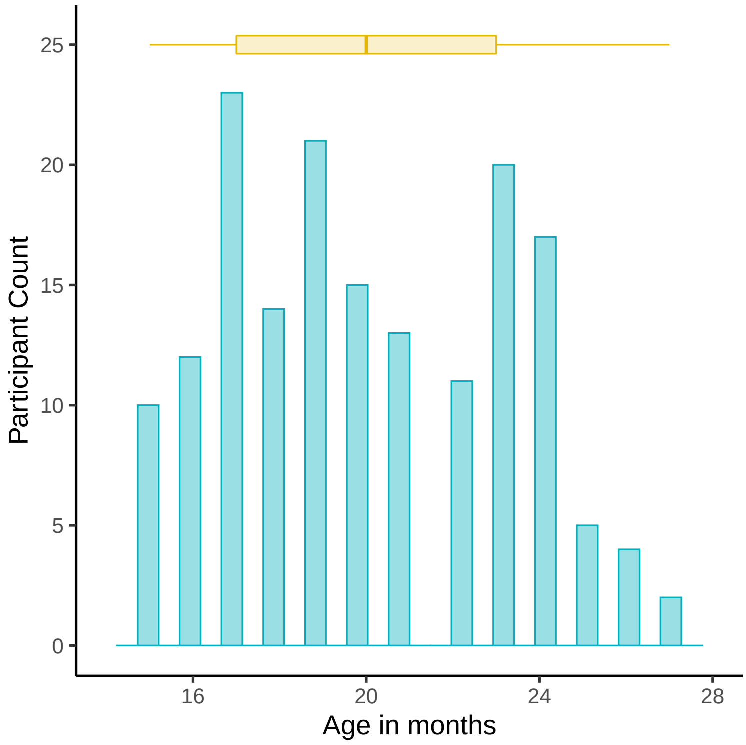


**Figure A.1:** Age distribution for the Test Group (approx. 18-months-old) on ELP UK

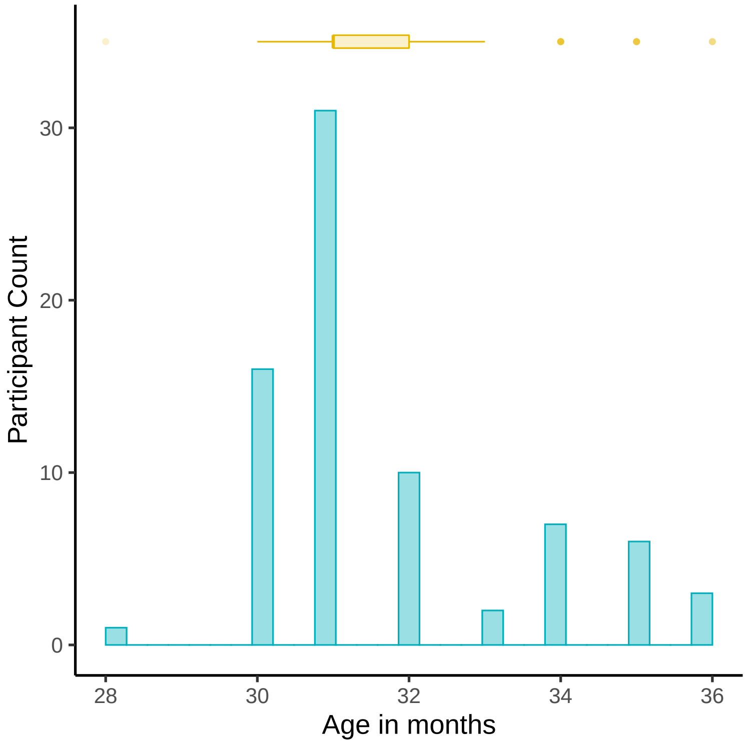


**Figure A.2:** Age distribution for the Retest Group (approx. 30-months-old) on ELP UK

**Table A.1:** Regression results for ELP Test at 18 months Age OCDI Comparisons (Model including Age)

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 191.42 | 1.00 | 0.00 | *** |
| ot1 | 29.82 | 1.00 | 0.00 | *** |
| ot2 | 7.32 | 1.00 | 0.01 | ** |
| ot3 | 10.91 | 1.00 | 0.00 | *** |
| Age | 3.77 | 1.00 | 0.05 | . |
| TrialType | 422.88 | 3.00 | 0.00 | *** |
| ot1:Age | 21.22 | 1.00 | 0.00 | *** |
| ot2:Age | 6.50 | 1.00 | 0.01 | * |
| ot3:Age | 15.01 | 1.00 | 0.00 | *** |
| ot1:TrialType | 163.41 | 3.00 | 0.00 | *** |
| ot2:TrialType | 133.20 | 3.00 | 0.00 | *** |
| ot3:TrialType | 173.16 | 3.00 | 0.00 | *** |
| Age:TrialType | 234.11 | 3.00 | 0.00 | *** |
| ot1:Age:TrialType | 95.29 | 3.00 | 0.00 | *** |
| ot2:Age:TrialType | 106.01 | 3.00 | 0.00 | *** |
| ot3:Age:TrialType | 143.14 | 3.00 | 0.00 | *** |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), Age in months and Trial Type (Reinforcement, Comprehension, Referent Selection and Retention). Blank indicates p >.05, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table A.2:** Regression results for ELP Test at 18 months Age OCDI Comparisons in the UK sample (Model Including OCDI Comprehension)

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 733.17 | 1.00 | 0.00 | *** |
| ot1 | 60.93 | 1.00 | 0.00 | *** |
| ot2 | 11.59 | 1.00 | 0.00 | *** |
| ot3 | 0.00 | 1.00 | 0.96 | |
| OCDIcomp | 2.45 | 1.00 | 0.12 | |
| TrialType | 444.54 | 3.00 | 0.00 | *** |
| ot1:OCDIcomp | 43.28 | 1.00 | 0.00 | *** |
| ot2:OCDIcomp | 8.03 | 1.00 | 0.01 | ** |
| ot3:OCDIcomp | 2.11 | 1.00 | 0.15 | |
| ot1:TrialType | 936.78 | 3.00 | 0.00 | *** |
| ot2:TrialType | 213.05 | 3.00 | 0.00 | *** |
| ot3:TrialType | 175.41 | 3.00 | 0.00 | *** |
| OCDIcomp:TrialType | 83.82 | 3.00 | 0.00 | *** |
| ot1:OCDIcomp:TrialType | 951.78 | 3.00 | 0.00 | *** |
| ot2:OCDIcomp:TrialType | 162.67 | 3.00 | 0.00 | *** |
| ot3:OCDIcomp:TrialType | 224.22 | 3.00 | 0.00 | *** |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), OCDI Comprehension Score and Trial Type (Reinforcement, Comprehension, Referent Selection and Retention). Blank indicates p >.05, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

**Table A.3:** Regression results for ELP Test at 18 months Age OCDI Comparisons in the UK sample (Model Including OCDI Production)

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 2407.98 | 1.00 | 0.00 | *** |
| ot1 | 68.64 | 1.00 | 0.00 | *** |
| ot2 | 6.82 | 1.00 | 0.01 | ** |
| ot3 | 0.00 | 1.00 | 0.98 | |
| OCDIprod | 3.07 | 1.00 | 0.08 | . |
| TrialType | 1391.25 | 3.00 | 0.00 | *** |
| ot1:OCDIprod | 58.49 | 1.00 | 0.00 | *** |
| ot2:OCDIprod | 7.07 | 1.00 | 0.01 | ** |
| ot3:OCDIprod | 15.18 | 1.00 | 0.00 | *** |
| ot1:TrialType | 873.65 | 3.00 | 0.00 | *** |
| ot2:TrialType | 231.38 | 3.00 | 0.00 | *** |
| ot3:TrialType | 166.43 | 3.00 | 0.00 | *** |
| OCDIprod:TrialType | 141.58 | 3.00 | 0.00 | *** |
| ot1:OCDIprod:TrialType | 723.37 | 3.00 | 0.00 | *** |
| ot2:OCDIprod:TrialType | 97.76 | 3.00 | 0.00 | *** |
| ot3:OCDIprod:TrialType | 182.37 | 3.00 | 0.00 | *** |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear), ot2 (quadratic) and ot3 (cubic), OCDI Production Score and Trial Type (Reinforcement, Comprehension, Referent Selection and Retention). Blank indicates p >.05, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

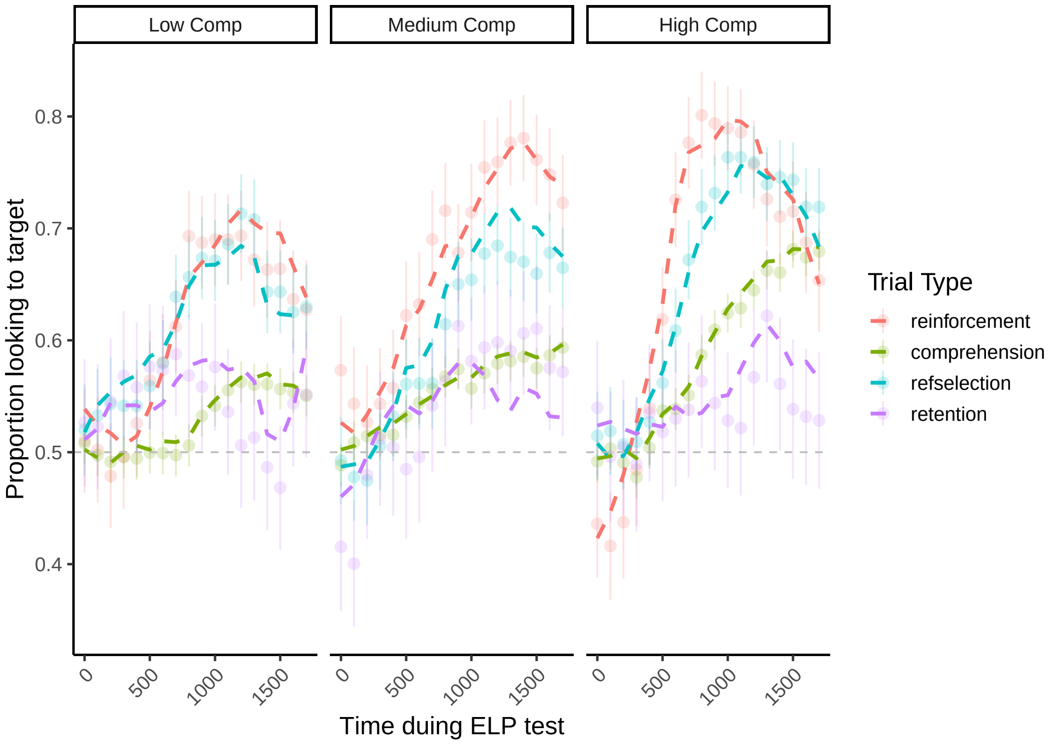


**Figure A.3:** Model predicted proportion looking to target on ELP trials from word onset using OCDI Comprehension instead of Age in months, split by Trial Type in the UK sample. Grey dashed line depicts chance performance (0.50). OCDI Comprehension is divided in tertiles for visualisation purposes. Points show the raw mean data per each 100 ms time bin with standard deviation. Dashed lines show the model predictions.

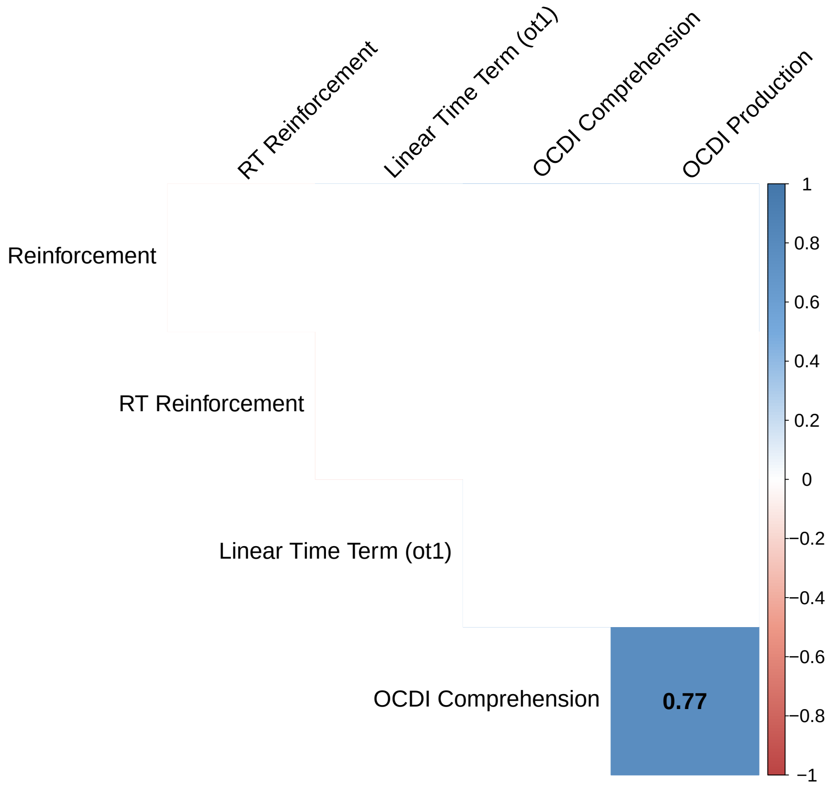


**Figure A.4:** Correlation matrix for the three ELP Reinforcement measures in the UK sample. Only correlations smaller than 0.01 (sig.level <0.01) are displayed. Correlations were performed only on those children who had data for all the subtasks in this section. Positive correlations are in blue and negative correlations in red, the strength of the colour indicates the strength of the relationship.

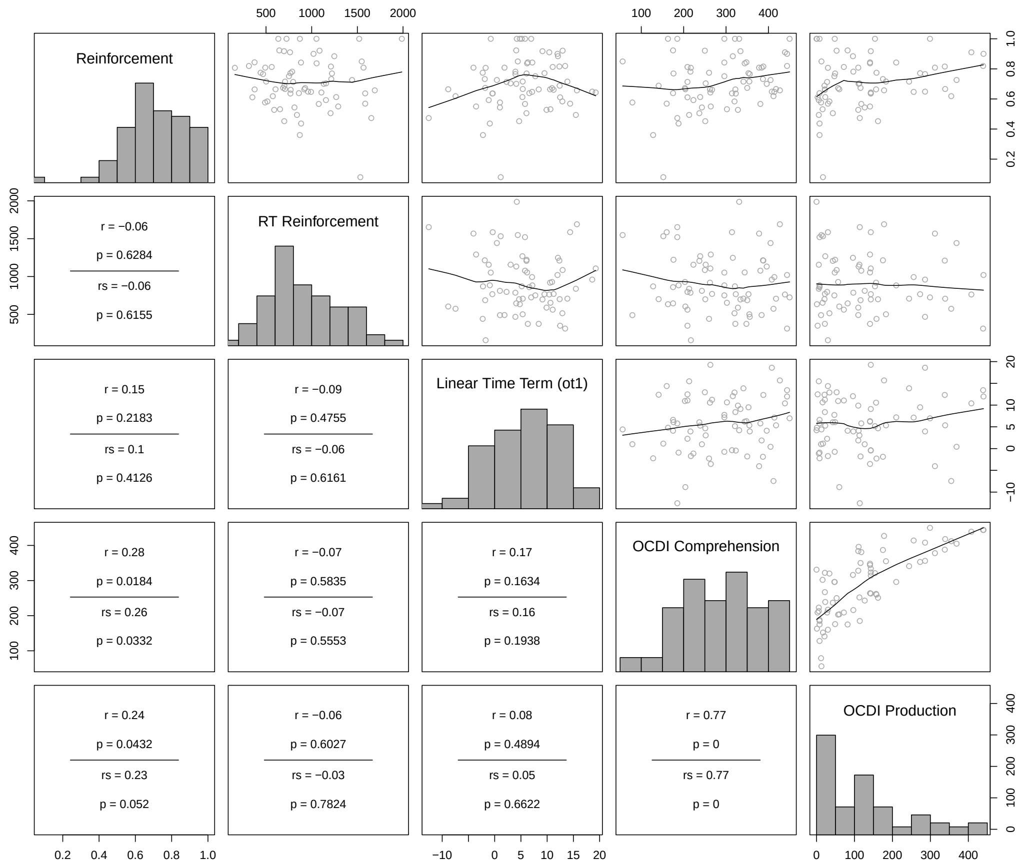


**Figure A.5:** Correlation matrix for the three ELP Reinforcement measures in the UK sample, including scatterplots with Pearson (r) and Spear- man (rs) correlations in the lower triangle with p-values. Smothers showing the relationships between variables are added to panels in the upper triangle, and histograms are added to the panels on the diagonal. Correlations were performed only on those children who had data for all the subtasks in this section.

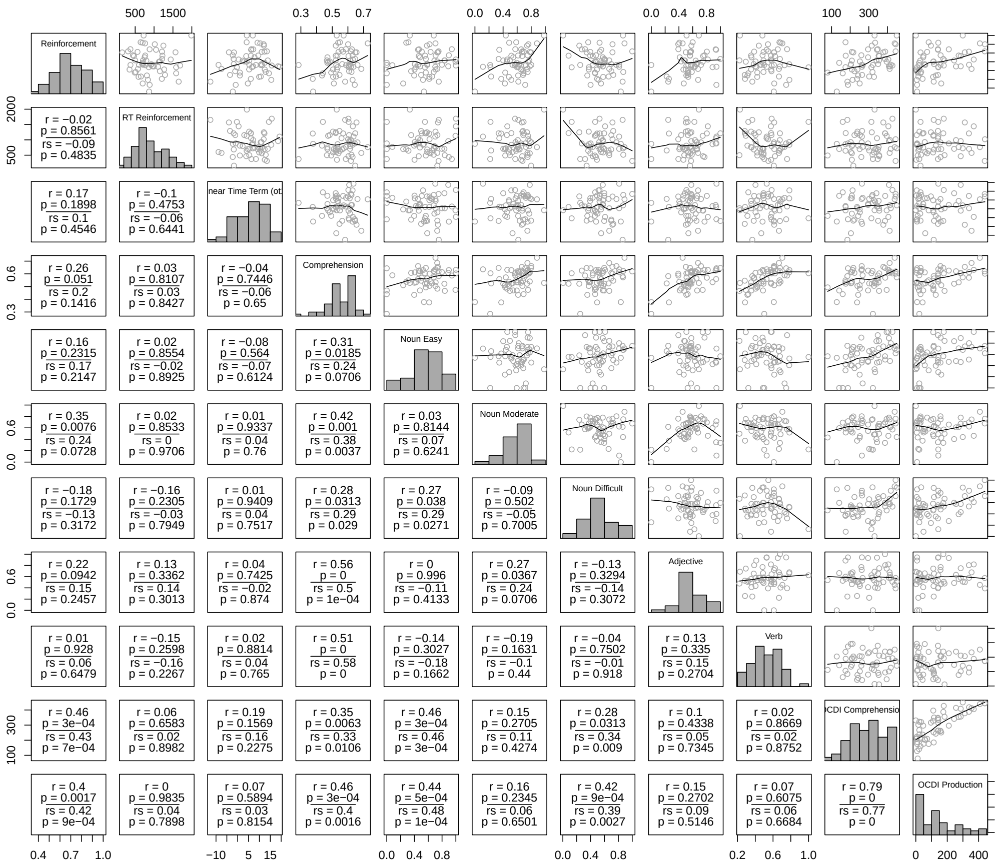

**Figure A.6:** Correlation matrix for ELP Reinforcement, ELP Comprehension and OCDI scores in the UK sample, including scatterplots with Pearson (r) and Spearman (rs) correlations in the lower triangle with p-values. Smothers showing the relationships between variables are added to panels in the upper triangle, and histograms are added to the panels on the diagonal. Correlations were performed only on those children who had data for all the subtasks in this section.

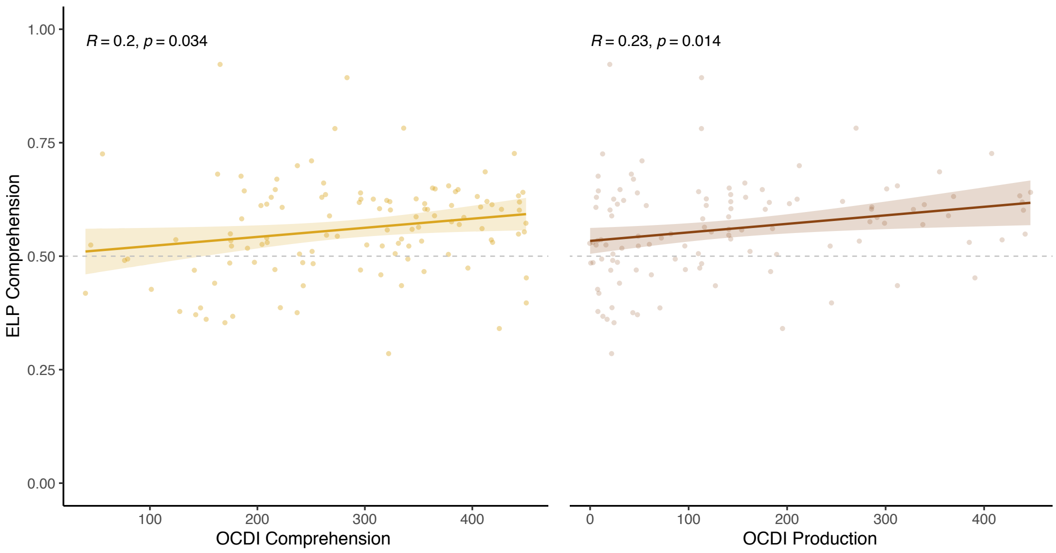


**Figure A.7:** Correlation between OCDI Comprehension and OCDI Production Total Scores and ELP Comprehension trials in the UK sample, computed as the mean proportion looking to target on the window of interest during test. Pearson correlation (R) and p-values are indicated.

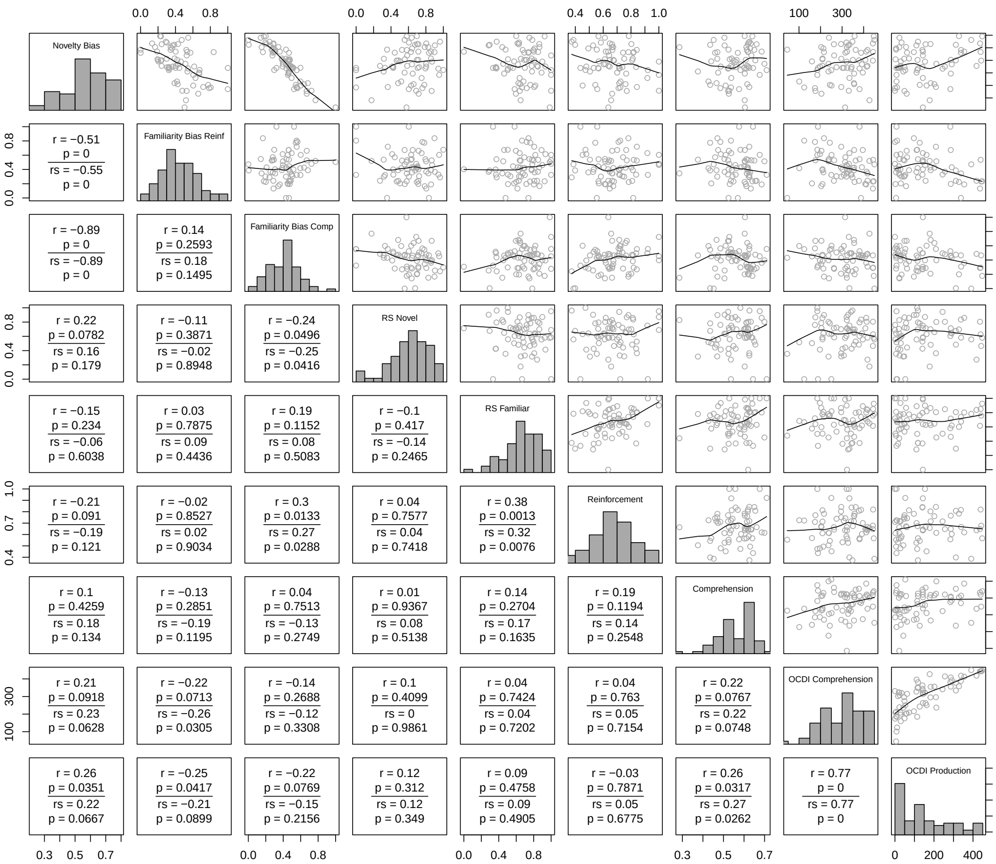


**Figure A.8:** Correlation matrix for Reinforcement, Comprehension and OCDI measures in the UK sample, including scatterplots with Pearson (r) and Spear- man (rs) correlations in the lower triangle with p-values. Smothers showing the relationships between variables are added to panels in the upper triangle, and histograms are added to the panels on the diagonal. Correlations were performed only on those children who had data for all the subtasks in this section.

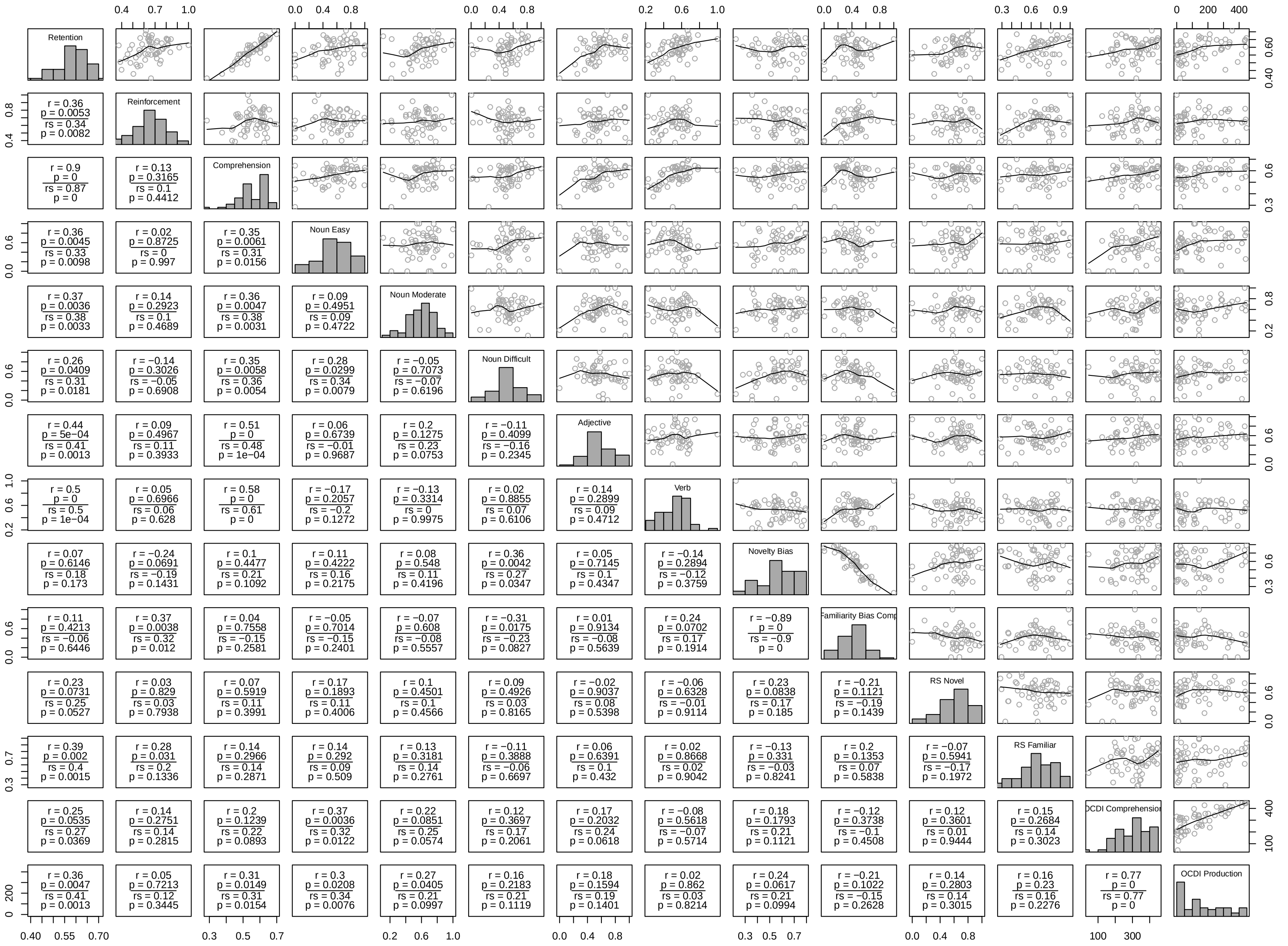


**Figure A.9:** Correlation matrix for Retention and other ELP and OCDI measures in the UK sample, including scatterplots with Pearson (r) and Spear- man (rs) correlations in the lower triangle with p-values. Smothers showing the relationships between variables are added to panels in the upper triangle, and histograms are added to the panels on the diagonal. Correlations were performed only on those children who had data for all the subtasks in this section.

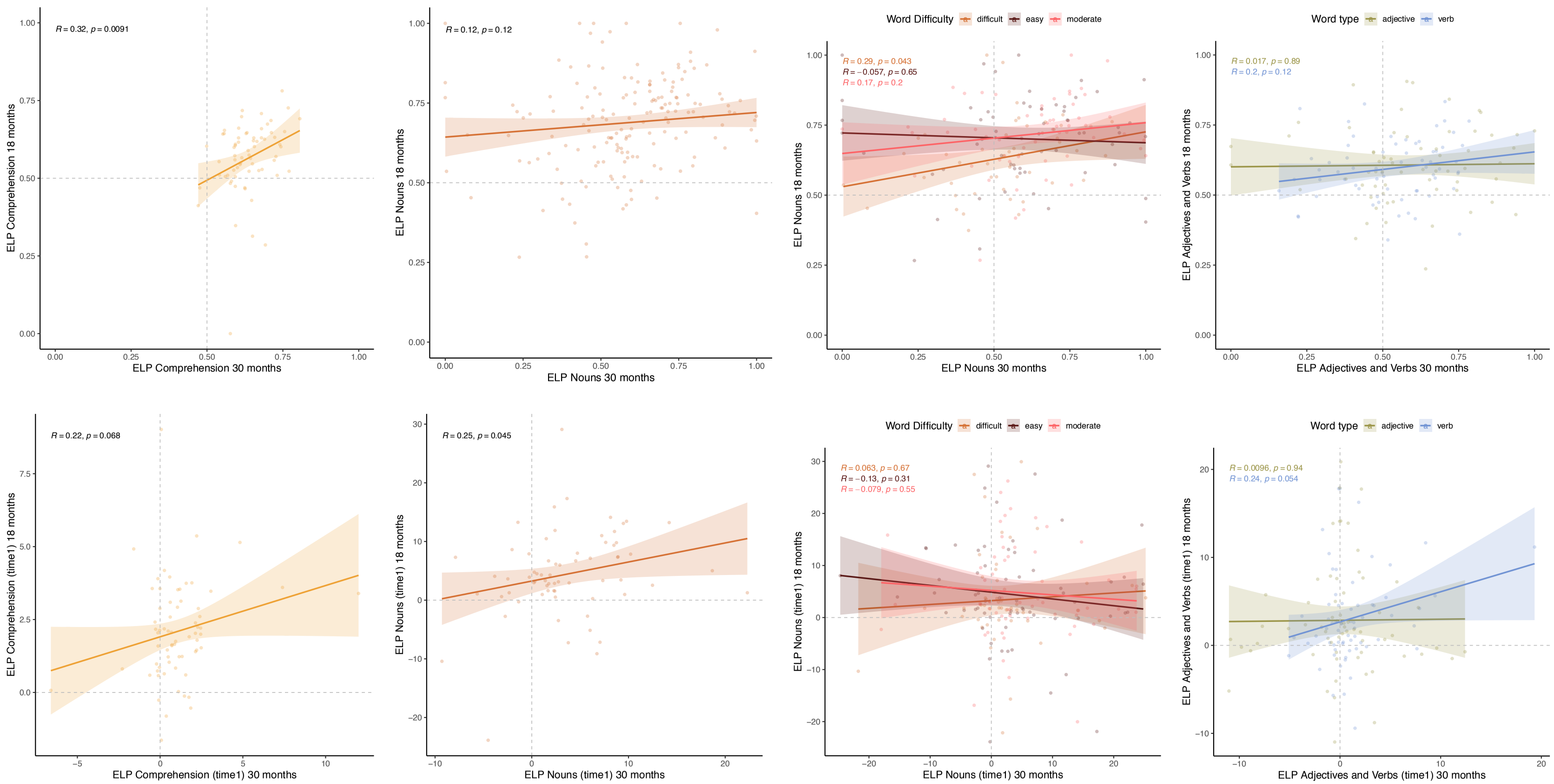


**Figure A.10:** Pearson correlations between ELP Comprehension at 18 months versus ELP Comprehension at 30 months in the UK sample. On the top row mean proportion of looks to target during the window of interest for overall Comprehension, Nouns, Nouns by difficulty, Adjectives and Verbs by test type (18 versus 30-months). On the bottom row model coefficients (time1) for overall Comprehension, Nouns, Nouns by difficulty, Adjectives and Verbs by test type (18 versus 30-months). Each point represents a participant. The regression line with standard deviation is fitted to the data for an easier visualisation of the relationship. Correlation (R) and p-value (p) are indicated at the top.

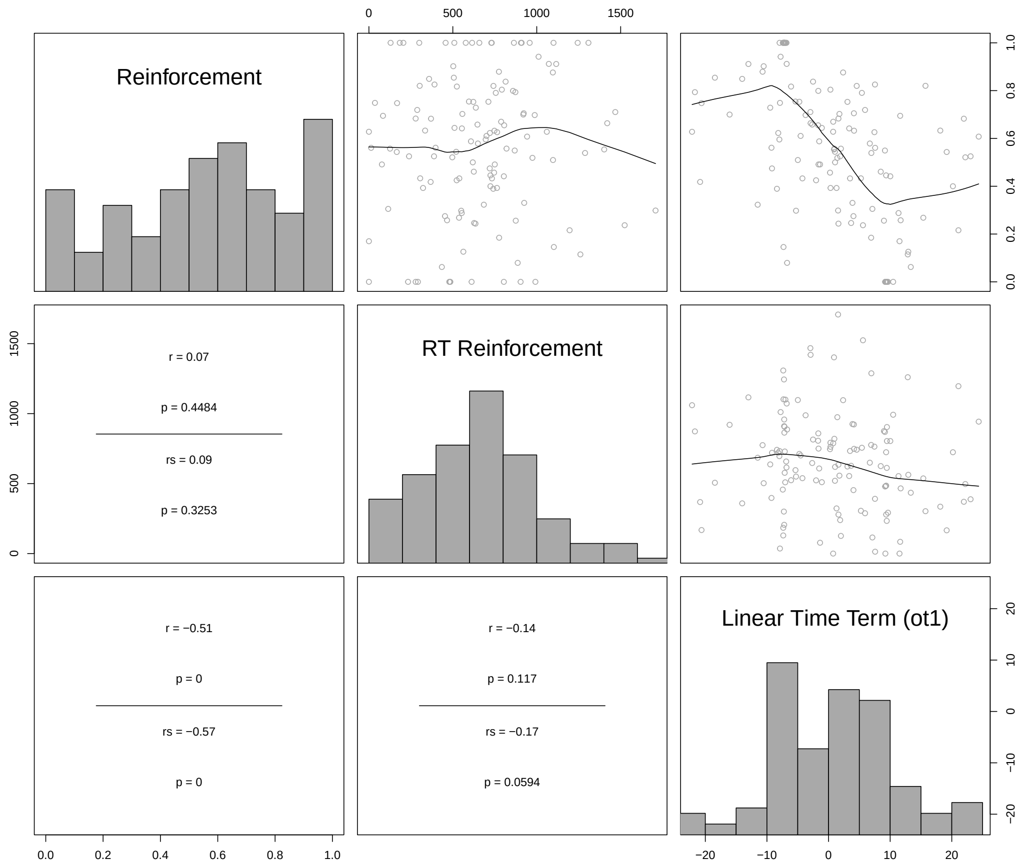


**Figure A.11:** Correlation matrix for the three ELP Reinforcement measures in the India sample, including scatterplots with Pearson (r) and Spear- man (rs) correlations in the lower triangle with p-values. Smothers showing the relationships between variables are added to panels in the upper triangle, and histograms are added to the panels on the diagonal. Correlations were performed only on those children who had data for all the subtasks in this section.

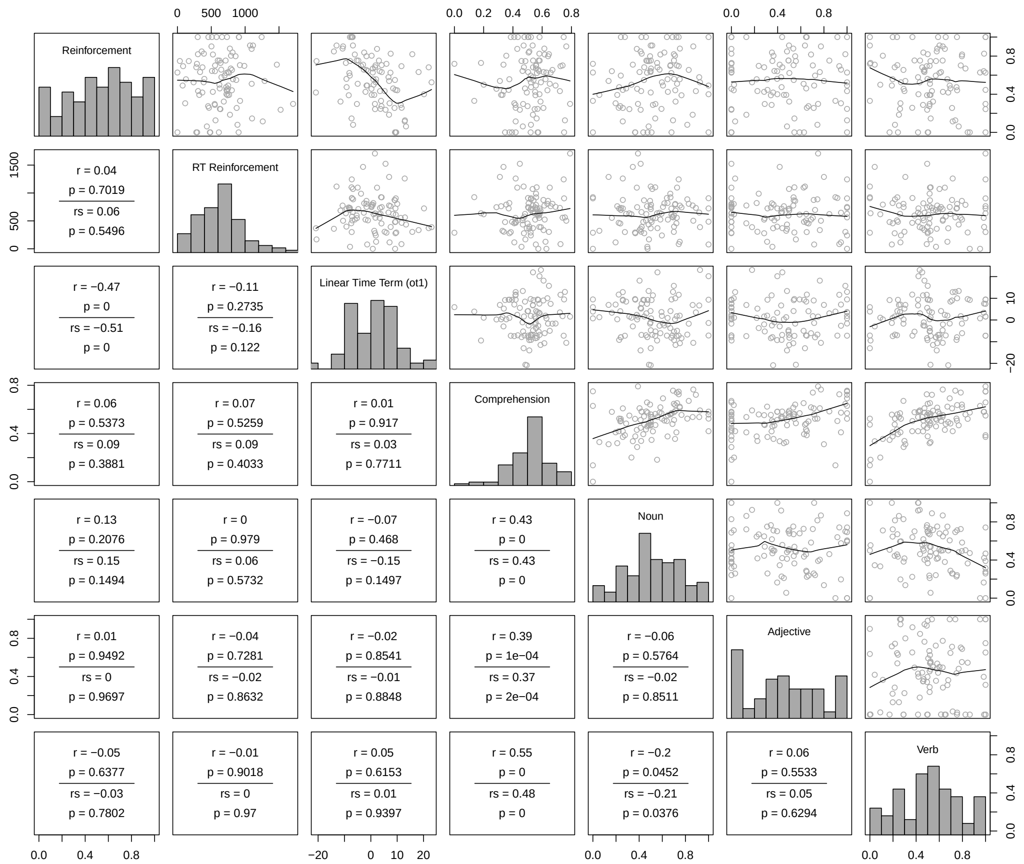

**Figure A.12:** Correlation matrix for ELP Reinforcement and ELP Comprehension measures in the India sample, including scatterplots with Pearson (r) and Spear- man (rs) correlations in the lower triangle with p-values. Smothers showing the relationships between variables are added to panels in the upper triangle, and histograms are added to the panels on the diagonal. Correlations were performed only on those children who had data for all the subtasks in this section.

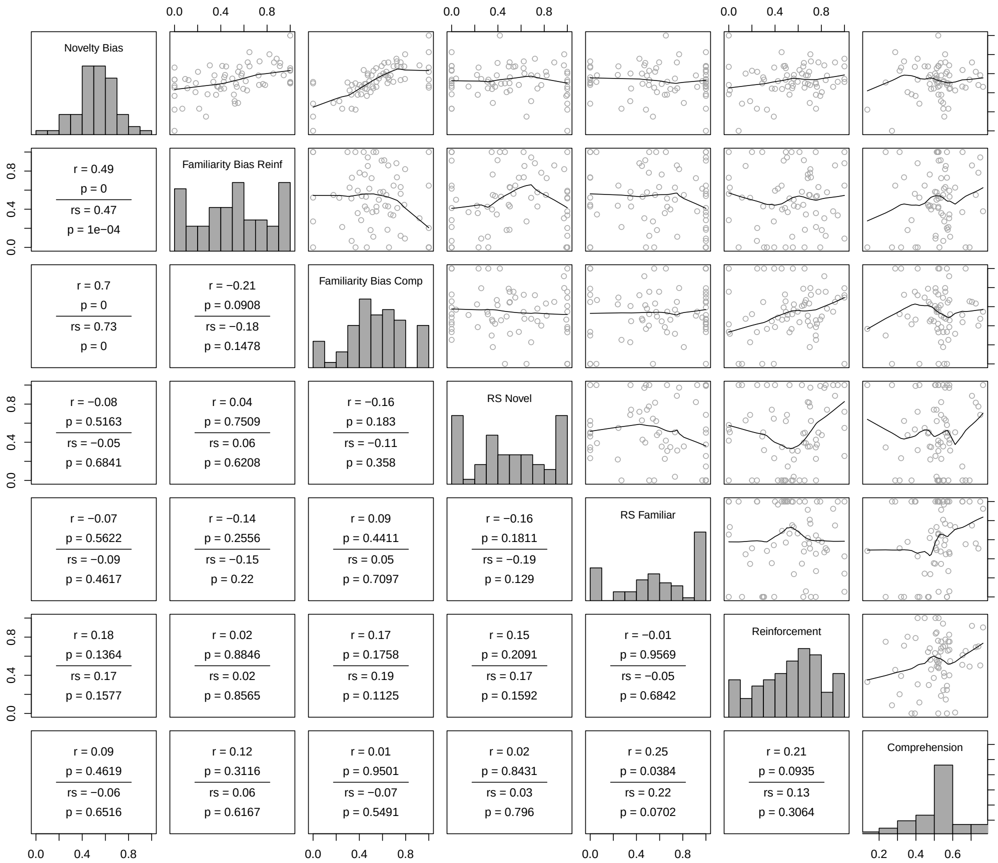


**Figure A.13:** Correlation matrix for ELP Reinforcement, Comprehension, Novelty and Familiarity biases and RS measures in the India sample, including scatterplots with Pearson (r) and Spear- man (rs) correlations in the lower triangle with p-values. Smothers showing the relationships between variables are added to panels in the upper triangle, and histograms are added to the panels on the diagonal. Correlations were performed only on those children who had data for all the subtasks in this section.

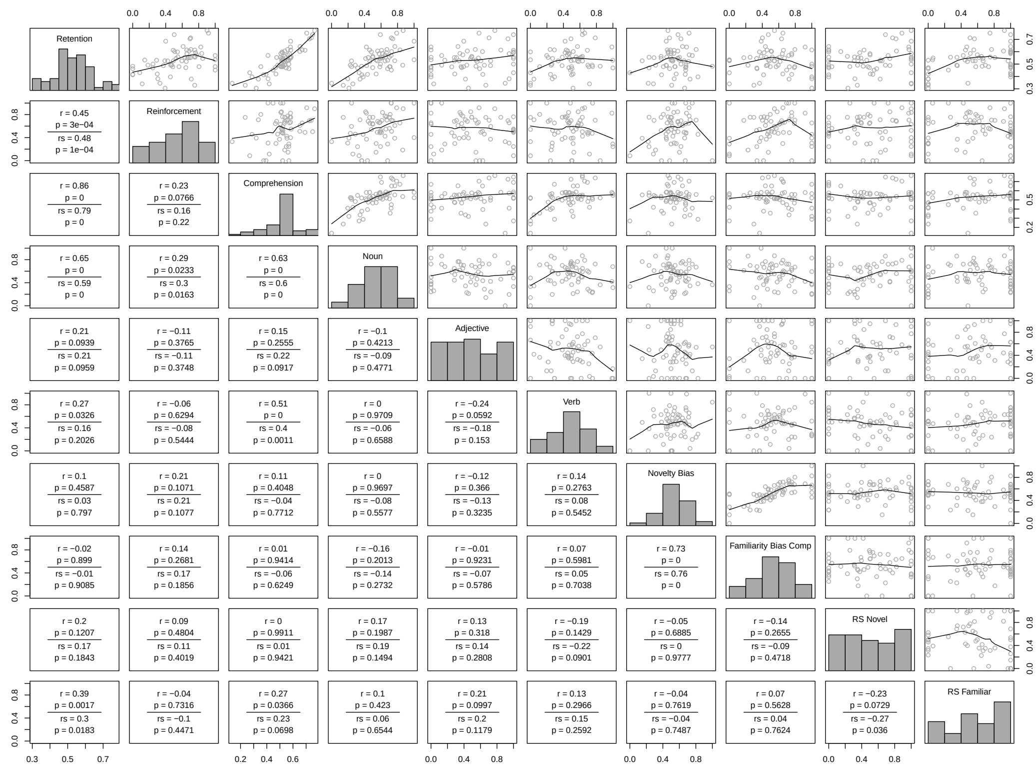

**Figure A.14:** Correlation matrix for ELP Retention and other ELP measures in the India sample, including scatterplots with Pearson (r) and Spear- man (rs) correlations in the lower triangle with p-values. Smothers showing the relationships between variables are added to panels in the upper triangle, and histograms are added to the panels on the diagonal. Correlations were performed only on those children who had data for all the subtasks in this section.

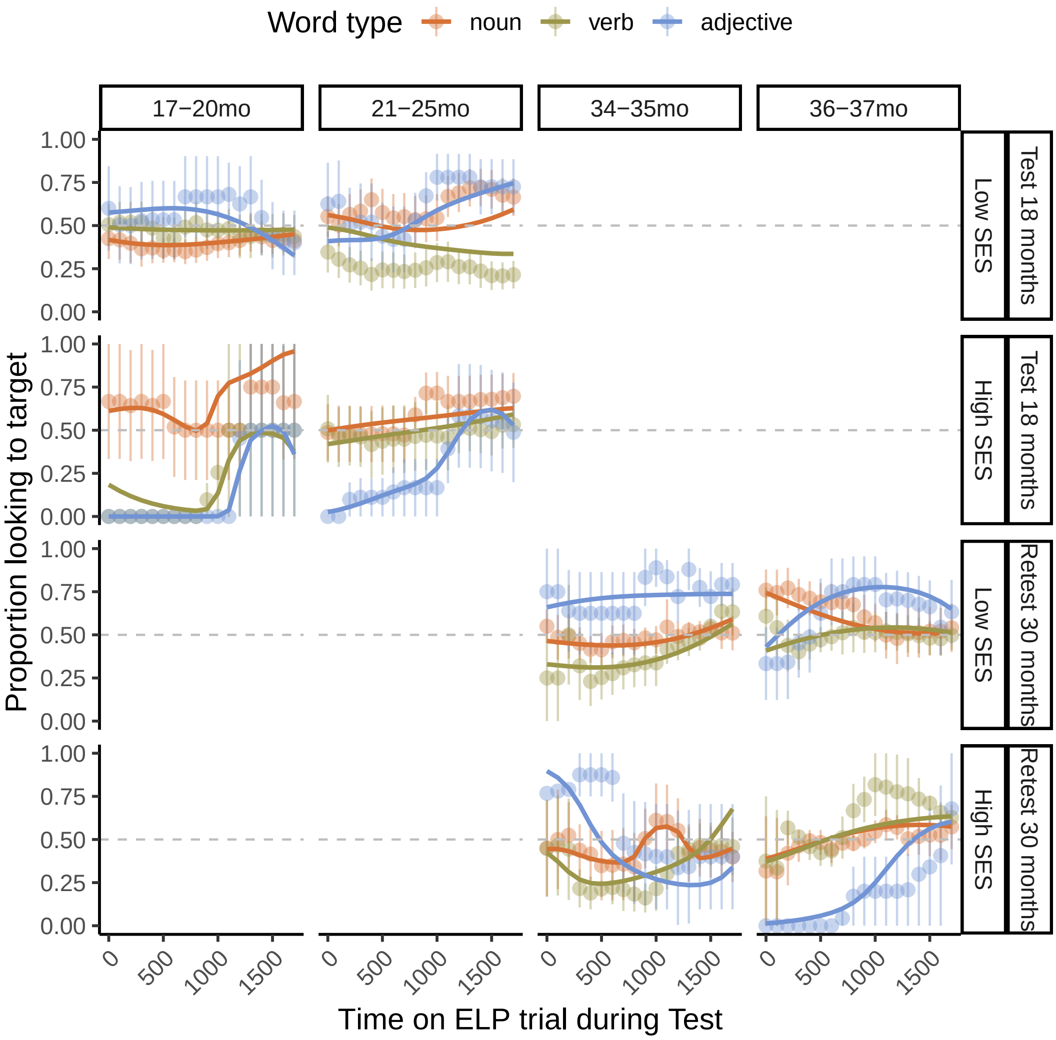


**Figure A.15:** Model predicted proportion looking to target in ELP India Comprehension trials split by Word Type (nouns, verbs and adjectives) by Age (in months), SES based on maternal education and Test Type (test at 18 months versus retest at 30 months). Grey dashed line depicts chance performance (0.50). Age in months is split in age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

**Table A.4:** Regression results for ELP India Comprehension Trials split by Word Type at 18 and 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 0.94 | 1.00 | 0.33 | |
| ot1 | 0.02 | 1.00 | 0.89 | |
| ot2 | 0.01 | 1.00 | 0.94 | |
| TestType | 1.05 | 1.00 | 0.30 | |
| Age | 0.72 | 1.00 | 0.40 | |
| WordType | 847.44 | 2.00 | 0.00 | *** |
| SES | 1.38 | 1.00 | 0.24 | |
| ot1:TestType | 0.64 | 1.00 | 0.42 | |
| ot2:TestType | 0.65 | 1.00 | 0.42 | |
| ot1:Age | 0.00 | 1.00 | 0.96 | |
| ot2:Age | 0.00 | 1.00 | 0.97 | |
| TestType:Age | 0.69 | 1.00 | 0.41 | |
| ot1:WordType | 120.80 | 2.00 | 0.00 | *** |
| ot2:WordType | 41.03 | 2.00 | 0.00 | *** |
| TestType:WordType | 168.94 | 2.00 | 0.00 | *** |
| Age:WordType | 811.90 | 2.00 | 0.00 | *** |
| ot1:SES | 0.00 | 1.00 | 0.95 | |
| ot2:SES | 0.03 | 1.00 | 0.86 | |
| TestType:SES | 0.61 | 1.00 | 0.43 | |
| Age:SES | 1.22 | 1.00 | 0.27 | |
| WordType:SES | 1102.04 | 2.00 | 0.00 | *** |
| ot1:TestType:Age | 0.55 | 1.00 | 0.46 | |
| ot2:TestType:Age | 0.55 | 1.00 | 0.46 | |
| ot1:TestType:WordType | 164.23 | 2.00 | 0.00 | *** |
| ot2:TestType:WordType | 111.66 | 2.00 | 0.00 | *** |
| ot1:Age:WordType | 117.74 | 2.00 | 0.00 | *** |
| ot2:Age:WordType | 40.67 | 2.00 | 0.00 | *** |
| TestType:Age:WordType | 204.42 | 2.00 | 0.00 | *** |
| ot1:TestType:SES | 0.59 | 1.00 | 0.44 | |
| ot2:TestType:SES | 1.08 | 1.00 | 0.30 | |
| ot1:Age:SES | 0.00 | 1.00 | 0.96 | |
| ot2:Age:SES | 0.02 | 1.00 | 0.89 | |
| TestType:Age:SES | 0.31 | 1.00 | 0.58 | |
| ot1:WordType:SES | 174.00 | 2.00 | 0.00 | *** |
| ot2:WordType:SES | 41.41 | 2.00 | 0.00 | *** |
| TestType:WordType:SES | 168.00 | 2.00 | 0.00 | *** |
| Age:WordType:SES | 1125.73 | 2.00 | 0.00 | *** |
| ot1:TestType:Age:WordType | 134.27 | 2.00 | 0.00 | *** |
| ot2:TestType:Age:WordType | 76.36 | 2.00 | 0.00 | *** |
| ot1:TestType:Age:SES | 0.54 | 1.00 | 0.46 | |
| ot2:TestType:Age:SES | 0.90 | 1.00 | 0.34 | |
| ot1:TestType:WordType:SES | 284.59 | 2.00 | 0.00 | *** |
| ot2:TestType:WordType:SES | 55.10 | 2.00 | 0.00 | *** |
| ot1:Age:WordType:SES | 181.11 | 2.00 | 0.00 | *** |
| ot2:Age:WordType:SES | 40.22 | 2.00 | 0.00 | *** |
| TestType:Age:WordType:SES | 338.11 | 2.00 | 0.00 | *** |
| ot1:TestType:Age:WordType:SES | 291.79 | 2.00 | 0.00 | *** |
| ot2:TestType:Age:WordType:SES | 50.48 | 2.00 | 0.00 | *** |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear) and ot2 (quadratic), Age in months, Word Type (nouns, verbs and adjectives) and SES based on maternal education. Blank indicates p >.05, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

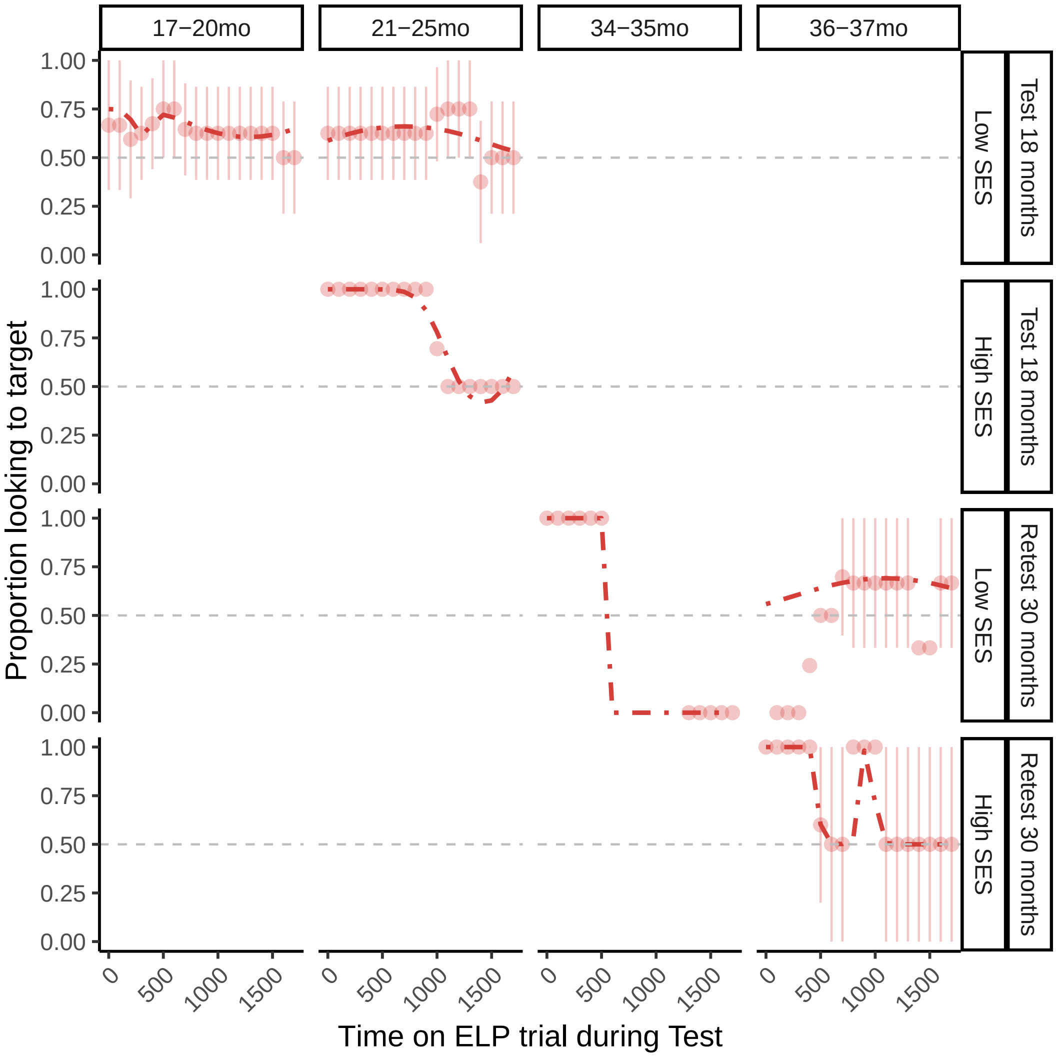


**Figure A.16:** Model predicted proportion looking to target in ELP India Retention trials by Age (in months), SES based on maternal education and Test Type (test at 18 months versus retest at 30 months). Grey dashed line depicts chance performance (0.50). Age in months is split in age groups to facilitate visualization. Points show the raw mean data per each 100 ms time bin with standard deviation. Line shows the model predictions.

**Table A.5:** Regression results for ELP India Retention Trials at 18 and 30 months

*Analysis of Deviance Table (Type III Wald chisquare tests)*

| term | statistic | df | p.value | significance |
|---|---|---|---|---|
| (Intercept) | 0.18 | 1.00 | 0.67 | |
| ot1 | 0.47 | 1.00 | 0.49 | |
| ot2 | 0.26 | 1.00 | 0.61 | |
| TestType | 3.61 | 1.00 | 0.06 | . |
| Age | 0.02 | 1.00 | 0.89 | |
| ot1:TestType | 9.92 | 1.00 | 0.00 | ** |
| ot2:TestType | 44.60 | 1.00 | 0.00 | *** |
| ot1:Age | 0.30 | 1.00 | 0.58 | |
| ot2:Age | 0.05 | 1.00 | 0.82 | |
| TestType:Age | 3.31 | 1.00 | 0.07 | . |
| ot1:TestType:Age | 9.40 | 1.00 | 0.00 | ** |
| ot2:TestType:Age | 54.81 | 1.00 | 0.00 | *** |

*Note. Fixed effects are displayed including the Time term represented as ot1 (linear) and ot2 (quadratic), Age in months, and SES based on maternal education. Blank indicates p >.05, * indicates p <.05, ** indicates p <.01, *** indicates p <.001*

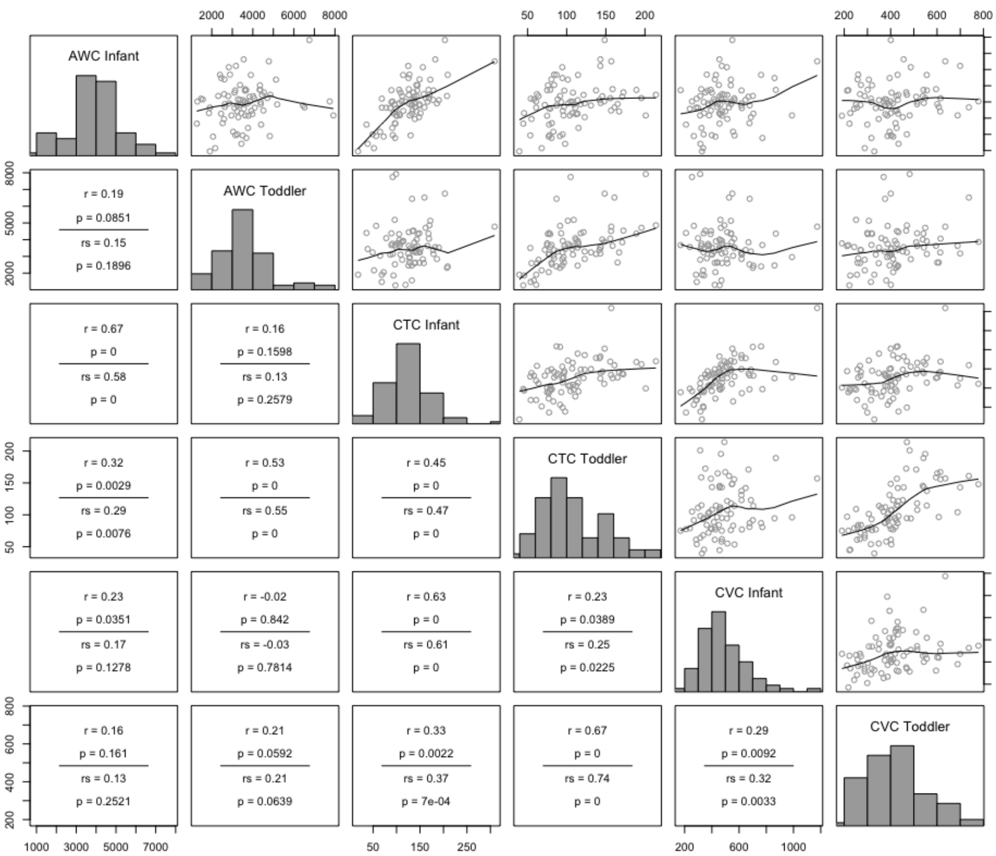


**Figure A.17:** Uncorrected matrix of scatterplots for LENA Infant and LENA Toddler in India including adult word count (AWC), turn count (TC) and child vocalisations (CVC) per hour. A matrix of scatterplots is produced with Pearson (r) and Spearman (rs) correlations in the lower triangle including p-values. Smoothers showing the relationships between variables are added to panels in the upper triangle, and histograms are added to the panels on the diagonal. Correlations were performed only on those children who had data for all the subtasks in this section.