Language Delay and Contextual Diversity: Individual and Environmental Differences in Early Word Learning

By

Eva María Jiménez Mesa

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

University of Warwick, Department of Psychology

September 2020

Contents

List of Tables

List of Figures

Acknowledgements

First of all, I am grateful to the Leverhulme Doctoral Bridges Program, which as well as providing valuable financial support during my PhD, also provided training in successfully bridging the boundaries of social sciences with mathematical sciences. Without the input from the people who participated in this programme, the output of my thesis would have been less fruitful.

I would like to express my sincere gratitude to all the families that participated in my main project. Despite the difficulties of participating remotely and over a long period of time, they remained enthusiastic about taking part. The main project of this thesis would not have been possible without the participation of these very generous families.

Enormous thank goes to my supervisor, Thomas Hills. He has been the most influential person I had during my PhD. His constant support and mentorship have been key in my journey towards becoming a successful researcher. I learned an immense amount from his thoughtful feedback on my work and enjoyed discussing ideas, problems and hypothesis with him. I am very fortunate to have Thomas as my supervisor.

I would also like to express my gratitude to Sotaro Kita, who was my second supervisor. He always showed interest in the progress of my projects. He provided valuable feedback and ideas, which I appreciated very much. I am also thankful that Kita will be my mentor during my upcoming Early Career Fellowship. I feel very confident that Kita will continue to prepare me for the academic world in the best possible way.

A special thanks to Eileen Haebig. I am lucky to have found a collaborator like Eileen. We share our enthusiasm for research on children with ASD. Her knowledge and constructive feedback on my writing have helped me become a better researcher.

I would like to thank Kate Messenger. She has been a great mentor during my PhD, guided me in the recruitment of participants as well as advising on making sure my main study complied with all ethical requirements.

I am also grateful to Bodo Winter, James Adelman and Massimo Stella for spending time guiding me on the statistical analysis of my data.

Also, I would like to thank the amazing people involved in the Warwick's baby lab for their immense contribution to recruit and maintain contact with the families in the lab's database. Especially I would like to thank Suzanne Aussems, for guiding me on the process of using the lab's database and her advice on how to approach parents. Without the baby lab's help, it would have been impossible to recruit as many families as I did for the main study.

I would like to thank the members of our lab's 'Pancake Gang': Tomas Engenthaler, Masita, Li Ying and Cynthia Siew. They have been a great source of support and friendship during the first years of my PhD. I missed them greatly this past year, but I am happy to see how each of them is fulfilling their dreams somewhere else in the world.

I appreciate the support of my parents. The following words are in Spanish for them: Quiero dar las gracias a mis padres Manuela y Juan Luís, que siempre me han dado su apoyo incondicional en todas las decisiones que he tomado en mi vida, incluido la de hacer este doctorado.

Finally, but not least, I would like to thank my husband, Keith. He always believed in my potential as a researcher and encouraged me to do a PhD. As well as supporting me emotionally during the ups and downs of the past four years, he also made a crucial contribution by creating the smartphone applications that allowed the remote data collection.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. Parts of this dissertation have been published or submitted to some journals by myself:

Chapter 2: Jiménez, E., Haebig, E., & Hills, T (under review). Identifying Areas of Overlap and Distinction in Early Lexical Profiles. *Journal of Autism and Developmental Disorders.*

Chapter 3: Jiménez, E., & Hills, T (under review). Semantic Maturation During the Comprehension-Expression Gap in Late and Typical Talkers. *Psychological Science.*

Chapter 4: Jiménez, E., & Hills, T (under review). The Richness of Child-Directed Speech is Reflected in Toddlers' Semantic Networks and Language Delay. *Child Development.*

Chapter 5: Jiménez, E., & Hills, T (under review). Untangling the Riddle of Contextual Diversity and Contextual Distinctiveness in Early Word Learning. *Child Development.*

My supervisor, Thomas Hills, supported my research throughout and provided feedback on the writing.

Summary

The production of the child's first words is the result of mapping the words heard to the correct referent in the world as well as learning the words' semantic characteristics. Experiencing the word in different linguistic contexts facilitates the accomplishment of these two learning tasks. The amount of different linguistic contexts associated with a word is known as the word's *contextual diversity*. Individual differences in exploiting contextual diversity might result in differences in the type of words that young children acquire. At the same time, environmental differences might affect the contextual diversity quality of the words in the speech that children hear. Altogether, the complexity of the interaction of internal and external factors related to contextual diversity might lead to differences in the child's language outcome. The general expectation is that poor ability coupled with a poor environment would be associated with early language delay. This thesis investigates this hypothesis by four means: 1) examining young children's lexical profiles, 2) understanding the role of contextual diversity in word learning relative to other contextual measures, 3) investigating the underlying mechanisms involved in word learning through contextual diversity, and 4) evaluating the semantic richness of child-directed speech of parents whose children are late talkers, typical talkers or late bloomers.

Chapter 1 introduces the topic and presents an outline of the thesis.

Chapter 2 investigates whether the lexical composition of late talkers and children with ASD, who also have language delay, differ from that of typical talkers. The most striking finding was that late talkers and children with ASD showed higher proportions of verbs than typical talkers, which might explain the weaker noun bias they presented. Most differences appear to reflect the extent of the language delay between the groups. However, children with ASD produced fewer high-social verbs than neurotypical children, a difference that might be associated with ASD features.

Chapter 3 explores the preverbal semantic maturation of words and its influence on early lexical learning. Study 1 compares the vocabularies of a large sample of late and typical talkers. Late talkers showed lower contextual diversity in their noun vocabularies but higher in their verb vocabularies. Study 2 explains these differences by comparing

three computational models finding that language delay is associated with lower semantic maturation of words.

Chapter 4 asks whether the structural quality of child-directed speech is associated with late language onset and whether its network-structure is reflected in the structural properties of the children's vocabularies. The study examines the vocabulary trajectories of 63 children as well the child-directed speech that they received at home during a 6-months period. The language environment experienced by the three groups of children with different lexical growth rates (i.e., typical talkers, late talkers, and late bloomers) differed in their structural-network properties, with the environment of late talkers being less semantically rich and less well-connected. The semantic associations between words that late talkers learned from their environments led to the structure of their expressive vocabularies to have a weaker network structure. Further, late talkers understood a different set of words which made their receptive lexical networks to be less well-connected. The results suggested an association between the quality of the language environment and the children's lexical abilities as well as the semantic network-structure of their vocabularies, which, as I discuss in Chapter 4, might have consequences in how children process language.

Chapter 5 investigates the role of two opposing but related contextual features in facilitating early word acquisition: contextual diversity and contextual distinctiveness. Hierarchical regression analysis confirmed the importance of both word features for word acquisition. In contrast to previous findings, words with an intermediate degree of contextual distinctiveness were found to be the earliest to be acquired. In addition, the best measure of contextual diversity considers the immediate context (words surrounding the word within speech), suggesting that the main role of contextual diversity is the semantic enrichment of the word's concept in the child's mind.

Finally, Chapter 6 discusses the theoretical and practical implications of the findings. Contextual diversity is an important word feature whose main role appears to be to aid a word's semantic maturation in the child's mental lexicon. The ability of parents to modulate the contextual diversity in their speech seems to be reflected in the structure of children's vocabularies and appears to be associated with early language delay in the case where the language input is semantically poor.

Preface

After decades of discussion on the Nature or Nurture debate, there is a consensus view that neither nature nor nurture alone is the answer to the question, but rather both nature *and* nurture. Understanding how these two interrelate is key in developmental science. I have been highly influenced by Elizabeth Bates' views of language, and far from assuming innateness, I rather see the human language capacity as a ground-breaking content that emanates from non-domain-specific brain functions. As Bates herself would put it "language could be viewed as a new machine constructed entirely out of old parts" (Bates & MacWhinney, 1989, p. 10), with language being a unique human behaviour that emerged from the interplay of the existing cognitive components to satisfy environmental demands. However, one must keep in mind that the complexity generated by this interaction makes it difficult, if not impossible, to determine a single cause for a particular language outcome.

In this thesis, early language delay is the main type of developmental outcome I focus on. In an attempt to understand the underpinnings of this type of linguistic developmental trajectory, I examine three related aspects across three studies: the characteristics of the language outcome, the internal cognitive processes and components for language acquisition, and the linguistic structure of the learning environment.

The study of early lexical profiles is of interest because it can provide clues about potential differences in how the underlying cognitive components operate and how the linguistic input influences word acquisition. The claim that late-talking toddlers might be just 'slow learners' implies that all those general components involved in language acquisition are delayed. At the same time, this also implies that late talkers' lexical structure should be similar to that of younger typical talkers with a similar vocabulary size. However, our assumption of a 'language device' made of general-purpose cognitive components might incline us to believe that these components have their own developmental constraints (e.g., time) and that minimal variations in just one of them might trigger an outcome like late language onset. This idea supports the view that sees the origin of language outcomes as developmental disassociations (Capute & Accardo, 1996). Therefore, qualitative differences in lexical profiles might serve as evidence against a general language delay (i.e., a delay in every component of the verbal

domain), and rather it is just one or more of these cognitive components that might be operating at a different level. In the latter case, it is important to identify whether different operational levels of a component are the result of normal changes in development (e.g., older children exposed to different objects/words), or whether it is the reflection of true atypical functioning (e.g., not presenting noun bias). Therefore, the identification of areas of overlap and distinction across the developing lexical profiles of children with and without language delay is a first step into finding the necessary mechanistic accounts that explain the areas of distinction. In all three studies I present in this thesis I examine the lexical structure of late-talking toddlers: in Chapter 2, I inspect the proportions of semantic categories and syntactic classes in the vocabularies of children with language delay, and in Chapters 3 and 4 I evaluate the properties of their lexical semantic networks. Some early language delays are associated with specific disorders such as autism. In Chapter 2, I investigate the potential associations of particular early lexical structure with social-cognitive components that are known to be specific to children with autism spectrum disorder. Providing evidence of the influence of a non-language-specific domain (such as the social domain) in the lexical profiles of children would further support the view of a language device made of non-specific components.

Many cognitive processes are implicated in early language acquisition. My particular contribution to the relevant literature concentrates on those mechanisms and cognitive components that are involved in learning the word meaning. To do this, I propose an alternative to an existing computational model in semantic networks, which, unlike the previous model, characterises the semantic maturation that words undergo during a preverbal comprehension phase. In Chapter 3, I present this model and evaluate its performance by linking changes in its components to particular vocabulary trajectories as I fit the model to observed data. In addition, in Chapter 5 I sought to clarify the role of contextual diversity in word acquisition by two means: first, I compared contextual diversity as a predictor of the order of word acquisition with contextual distinctiveness, another contextual feature known to facilitate word learning; and second, I investigated the size of the context (nearby words versus entire documents) that best predict word order and explore its relation with the type of learning facilitative effect.

Cognitive components are not the only responsible factors for the variations in language development. To get the whole picture, the learning environment has to come

into play. Continuing with my interest in the word-meaning acquisition, in Chapter 4, I also evaluate the semantic richness in the parental speech of families of children with and without early language delay. The longitudinal characteristic of this study (whose data collection occurred over the majority of my PhD) had the particular advantage that some of the participants overcame the apparent language delay during the study, making them 'late bloomers'. This population is of immense interest since this investigation helps answer two important questions: what late talkers and late bloomers have in common that could be related to their late language onset, and what is different between them that could explain the late bloomers' rapid recovery. Answers to these questions not only provide evidence for how different language profiles emerge but also serve to inform in the design of palliative and potentially preventive language interventions.

Chapter 1: Introduction

Typically, children are expected to utter their first words around 12 months of age. Although this and other language milestones vary, children at the lowest end of the language development spectrum are frequently of concern. In an effort to determine the causes of early language delay, researchers have investigated internal and external factors that might contribute to the late-onset. Learning the meaning of words is primordial in language development. Infants do not simply map the word to the correct referent in the environment, but also learn characteristics about the referent. This thesis focuses on understanding the relation between the learning of the word's meaning and early language delay. This section starts by introducing the characteristics of the child's first words in relation to meaning and cognition, and describes how the language environment can be exploited by children to construct the words' concept. In particular, *contextual diversity* is introduced as an important word feature that facilitates word learning. Following this, a working definition of *late talker* is provided as well as a description and evaluation of the parental checklist used across all four studies of this thesis. Then, relevant literature is reviewed to outline the main internal and external factors that are related to language delay in late talkers. Importantly, hypotheses for potential relations between contextual diversity and language delay are defined, a methodological approach to defining the comparison group is described, and its use justified to explore the current research questions. Finally, the aims of the present thesis are defined, and the four studies are outlined.

1.1 The First Words

The first signs of language understanding in infants commence at very early stages of development, as early as six months of age (Bergelson & Swingley, 2012). It is a remarkable moment for a baby when she makes the first connection between what she hears and the corresponding event or object in the environment. From then on, the child starts what seems to be an effortless journey of learning the mother tongue over the next few years. After a few months of relatively clear understanding of many words and lots of babbling, the first spoken words emerge. What is peculiar about these first

words is that the infant uses them in different meaning contexts, signifying that she does not have a unique meaning linked to the word, but potentially a more general one. Lewis (1967) illustrated well how the word "mama" initially had a manipulative effect, uttered by the infant when he is in a moment of need, requesting the caregiver's attention. Later on, the same child in this example showed a declarative use of the word "mama". For instance, the same child showed contentment while saying the word "mama" in the mother's arms; and on another occasion, this child handed over his toys to a new visitor one by one, and uttered "mama" each time. This more general use of the word "mama" by the child is then naturally narrowed thanks to the caregivers' feedback to the child's utterance with gazes, gestures, movements towards objects, and more importantly, with their speech.

This general-to-specific pattern of development in the words' meaning is also observed in the first semantic categories that infants form. As early as two months old infants are able to categorise living and non-living things (Quinn & Johnson, 2000), and one or two months later they are capable of making distinctions between two more specific concepts such as "cat" and "dog" (Eimas & Quinn, 1994). Given that the first words that a child produces occur around the first year of life, it is not surprising that most of the first nouns already refer to basic-level concepts instead of superordinate concepts ("dog", and not "animal/living thing"). In fact, basic-level words represent approximately half of the noun types in child-directed speech (Feijoo et al., 2017), suggesting a positive influence of the language environment on the type of words that young children acquire first. The distinctive properties of most basic-level word categories (e.g., "dog") demonstrate their superiority over other higher ("animal") and lower ("Dalmatian") level categories (Rosch et al., 1976). For example, it is more difficult to distinguish between two types of dogs than between a dog and a fish because the perceptual similarity between dogs and the way we interact with dogs (motor similarity) is greater than those between a dog and a fish. Basic-level words are efficient when it comes to organising knowledge: they are informative enough about the category, with not too many features needed to define it. At the same time, basic-level words can be grouped in an adequate number of categories, allowing us to economise our limited cognitive resources (Collins & Quillian, 1969). All these characteristics offer the cognitive advantages of faster word retrieval and recognition (Rosch et al., 1976). Although early vocabularies include a large proportion of basic-level words, infants learn other types of words which are essential for adequate language development, such

as verbs, which are needed not just to express events, states or processes, but also to acquire grammar.

A general conception is that the first categorisation that young children make are based on salient perceptual features. This idea is supported by evidence that shows that the first concepts that children are able to discriminate are at a superordinate level (Quinn & Johnson, 2000) where less discrimination is required (Antonucci & Alt, 2011). For instance, children only need to be aware that people and animals can move around and have eyes which highly contrast with the non-living objects' features of stillness and facelessness. Later, to learn the category of "dog", the child has matured her discrimination abilities to notice that dogs are different from birds in that, among other features, dogs have legs and walk whereas birds have wings and fly. The sharp distinctiveness in perceptual characteristics is also evident in the children's first utterances. Age of word acquisition was predicted by object dissimilarity, with first words being more semantically dissimilar from other words than later acquired words (Engelthaler & Hills, 2017). Word distinctiveness was particularly strong for visual forms and surface features. This finding fits well with the common word learning strategy in infants of extending known noun words to new instances that are similar in shape (i.e., shape bias; Landau, Smith, & Jones, 1988).

However, children do not only categorise objects by similarity in perceptual features. In a clever experiment, 2-to-4-year-old children used causal information to determine whether an object was or was not a "blicket" (Gopnik & Sobel, 2000). Children placed the novel object in a "blicket detector" that would go off when a "blicket" was identified by the machine. The target "blickets" were purposely selected so that their shape and colour mismatched. Despite the visual incongruity, children were able to disregard the perceptual information (shape and colour) and instead used the causal power they experienced to classify objects as "blickets". Children not only can draw arbitrary associations between words and referents but also between a word and other words. For example, when young children start counting up to three, they probably associate these number words with each other because they usually appear together in parent's speech when an action is about to happen ("one, two, three, go!"), or when the adults point to individual objects sequentially within a set. These associations are not as obvious as perceptual features, but their frequent co-occurrence in child-direct speech make then salient and learnable. In contrast to using discrimination abilities for learning associations, the acquisition of associations like

"one, two, three" can be passively acquired from the learning environment. This last instance suggests two things: first, that for learning some word associations children depend on the frequency in which they co-occur in child-directed speech, making the language environment an important source for learning word associates; and second, that words encoded together (i.e. with their contextual companions) inevitably become part of each other's meaning.

1.2 The Context Matters

In ancient Roman and Greek times, the power of the context was used to win memory contests. With the *loci method*, the contestants memorised long lists of elements by organising them in association with mental images of physical places, such as the parts of a palace. The artificial associations between elements using two different modal entities helped them to activate and retrieve the requested item by mentally 'walking' to the physical place where the item was 'stored'. Similarly, people create associations between words early on in life. This is what happens when words tend to co-occur within speech, like "baby-bottle", "dog-cat" or "milk-cow". These pairs of words are also known as *free association norms* (Nelson, McEvoy, & Schreiber, 2004). The association between words is typically estimated by providing a cue word to the participant and then instructing them to give the first word that comes to mind (the target word). Free word association is a good predictor of lexical retrieval and language processing (Nelson, Schreiber, & McEvoy, 1992; Steyvers, Shiffirin, & Nelson, 2005). Interestingly, words that have been produced as a target in response to many cues are produced earlier by children (Hills, Maouene, Maouene, Sheya, & Smith, 2009a; Hills et al., 2010; Steyvers & Tenenbaum, 2005).

Some other times, semantic cues come from visual modalities, such as gestures. When learning novel words, two-year-old children who heard the novel word together with a gesture reinforcing either its function or its shape, showed superior fast- and slow-mapping than those children who were exposed to novel words alone (Capone & McGregor, 2005). Evidence demonstrated that practically anything that is in the environment during the encoding phase could be stored along with the word itself. For example, in adult bilingual people, memories in which the language used at the encoding context is the same as in the retrieval context are more detailed and emotional than memories in which the language of encoding and the retrieval context are incongruent (Larsen et al., 2002, Marian & Neisser, 2000, Matsumoto & Stanny, 2006). In a novel

word generalisation task, 2.5-to-3-year-old children performed better when the patterned cloth in the background at the testing phase matched with that of the learning phase (Vlach & Sandhofer, 2011). These studies provide evidence that words are not encoded alone but are richly accompanied by environmental elements that influence the way we organise and represent concepts in our minds.

1.3 What the Linguistic Structure Can Tell us about the Meaning

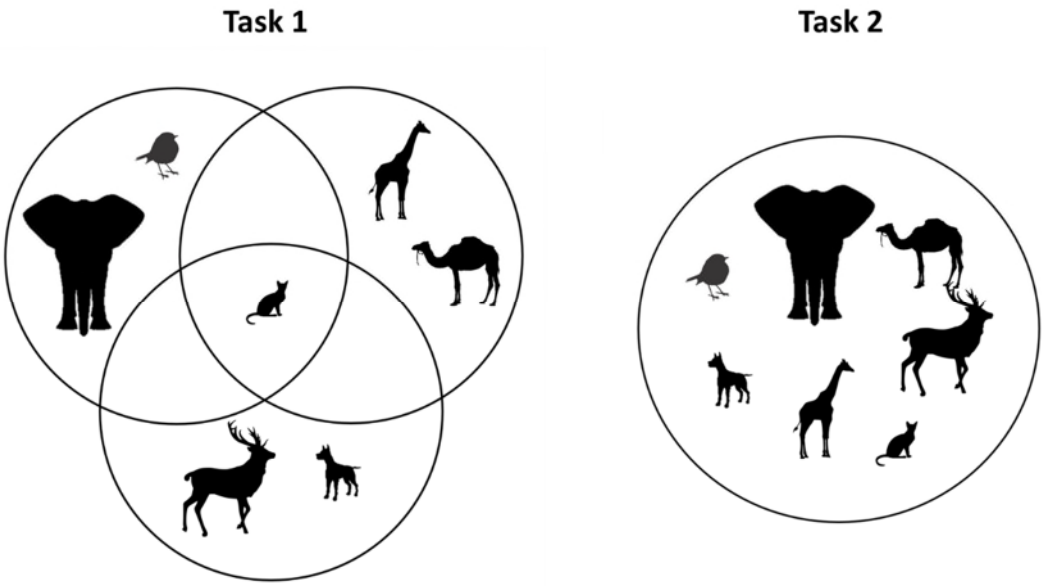
The focus of the present thesis is the semantic enrichment contained in the linguistic structure of parental speech, particularly that part of the meaning that can be abstracted from the interrelatedness between words. For example, the word 'wet' has special relations with words like 'towel', 'water', and 'rain', with which it co-occurs in speech. This distributional characteristic inherent in language amplifies and enriches the holistic meaning of the word 'wet' by adding new semantic nuances and connotations. Recent evidence based on eye-tracking data suggested that children perceive word pairs that frequently co-occur as being more semantically related than word pairs that share taxonomical features (Unger, Vales, & Fisher, 2020). By measuring word co-occurrence, one finds that some words are surrounded by a large number of diverse words, and some other words are always accompanied by a reduced number of words. This is the *contextual diversity* measure of words, typically computed as the number of different contexts in which the word appears (Adelman, Brown, & Quesada, 2006; Hills et al., 2010; Jones, Johns, Recchia, & Michael, 2012).

The role of contextual diversity in the cognitive system has been repeatedly investigated and compared to the role of word frequency. Word frequency has well-known positive effects as a facilitator for lexical access (Forster & Chambers, 1973; Krueger, 1975; Blythe, Liversedge, Joseph, White, & Rayner, 2009; Joseph, Nation, & Liversedge, 2013). The theoretical mechanisms behind this lexical advantage are that every encounter of the word reinforces its memory trace: this is the *principle of repetition*. However, in the last two decades, the focus of research interest has shifted onto contextual diversity, which is changing the views of how the lexicon is organised and accessed. The cognitive potential of contextual diversity is explained by the *principle of likely need*. The principle, foundational to many classic rational models of memory (Anderson & Milson, 1989; Anderson & Schooler, 1991; Dennis & Humphreys, 2001)

asserts that if a word has been encountered repeatedly during the learning phase, it has a high probability of being encountered in upcoming contexts, which intuitively requires a lexicon organised in a way that the word is easily accessible. Several studies have shown that words that have been learned in many different contexts are processed more easily and have the advantage of being recognised faster and more efficiently (Goldinger & Azuma, 2004; McDonald & Shillcock, 2001; Nelson & Shiffrin, 2006; Pexman et al., 2008; Frances, Martin, & Duñabeitia, 2020). Further, evidence of the superiority of contextual diversity over frequency was found in predicting reaction times in word naming and lexical decisions tasks (Adelman, Brown, & Quesada, 2006).

What does contextual diversity mean for an infant learning the maternal language? Typically, when children experience the same words and objects co-occurring across situations, they use these regularities to infer the words' referent. This is known as *cross-situational learning* (Siskind, 1996; Akhtar & Montague, 1999). However, the reduction in ambiguity is not always guaranteed as, even when the word appears across many situations, the words/objects that appear in those situations are the same, i.e. the *same* context. Words appearing in unique contexts make it difficult to solve the *referential of indeterminacy problem* (Quine, 1960). Therefore, the most straightforward positive effect of encountering the same word in *different* contexts is that it reduces referential ambiguity. Figure 1.1 illustrates the learning facilitation effect of contextual diversity. Each circle represents a context, and in each context, the child hears the novel word "dax". In Task 1, the high contextual diversity of the referent "cat" lowers the ambiguity of the task, allowing the child to map the novel word to the correct referent. In contrast, learning the meaning of the novel word "dax" in Task 2 seems impossible without any other cues.

Figure 1.1. Illustration of Learning a Word with High Contextual Diversity versus Learning a Word with Low Contextual Diversity.



Note. Each circle represents a unique context. The referent to the word "dax" can be deduced in Task 1 after the child hears the word three times, each time in a different context. In Task 2, hearing the word "dax" three times does not help to infer the correct referent of the word.

A second potential advantage of contextual diversity is that it might semantically enrich the words' mental representation. The semantic relatedness between words naturally found in language is reflected in how people mentally index concepts, as suggested by the high correlation between the words' contextual diversity and their number of associates in a free association task (Hills et al., 2010). The semantic enrichment that words acquire through contextual diversity has been exposed in studies that examined the electrophysiological correlates in the brain. Analysis of the event-related potentials (ERPs) has found larger N400 (a component related to lexical-semantic processing) for words with many associates (e.g., Müller, Duñabeitia, & Carreiras, 2010), many semantic features (e.g., Amsel, 2011), and for concrete words compared to abstract words (e.g. Barber et al., 2013). The larger amplitude of N400 in all these studies suggests that the process of retrieving a well-connected word is costlier (Holcomb, Grainger, & O'Rourke, 2002) supposedly due to the activation of the word's associates. Even though contextual diversity is highly correlated with word frequency,

Vergara-Martínez, Comesaña, and Perea (2017) demonstrated that the pattern in the amplitude of N400 is different for high-frequent words than for high contextual diversity words. As Vergara-Martínez and colleagues indicate, this finding reveals a lexical retrieval mechanism related with the activation of the semantic features (contextual diversity/word associates), which differs from mechanisms that deal with the strength of memory traces (word frequency).

1.4 Contextual Diversity as a Potential Factor Influencing Language Outcomes

The previous sections have described how the first words that children produce do not have a specific meaning, unlike the case of adults. However, contextual cues provided by caregivers are of importance to delimit the meaning of words. These early words have important cognitive advantages and are frequent in the linguistic input the child receives. However, children are active learners and do not acquire everything they hear. In fact, much of the language they learn is dependent on their perceptual discrimination abilities, which explain the characteristic semantic distinctiveness of their first words. Children also use their abilities to exploit the natural structure embedded in parental speech. In particular, contextual diversity has been shown to have a significant role in enriching the mental concept attached to the word, and children seem to be able to utilise contextual diversity cues in the language environment from early on in life. Despite the importance of contextual diversity in language acquisition, very little is known about how it exactly influences language development in early years. This raises the question of whether a failure to exploit contextual diversity as a learning mechanism might give rise to atypical language outcomes. Further, word learning through contextual diversity is not only dependent on the child's abilities to extract this information from the environment. The structure of the linguistic context must allow the child access to the different semantic nuances of the words through their co-occurrence with other words. This also raises the question of whether parental speech poor in contextual diversity (i.e., using words in unique, consistent linguistic contexts) has a negative impact on young children's language outcomes. In the next sections, I provide a working definition of *late talker* and review some relevant studies that focus on internal and external variables related to language delay in early childhood.

1.5 Early Language Delay

Late talkers are commonly identified at the age of two, with a prevalence of 10-20% (Desmarais et al., 2008, Klee et al., 1998, Reilly et al., 2007). While many late talkers resolve by school age (Paul, 1991; Rescorla, Roberts, & Dahlsgaard, 1997; Feldman et al., 2005), some remain delayed and are later diagnosed with developmental language disorder (DLD, Bishop et al., 2012; Taylor et al., 2013). In many cases, early language delay is caused by physical (e.g., ear infection), environmental (e.g., multilingual environment), or neurological problems (e.g., Down syndrome or Autism). However, some late-talking toddlers do not present any primary condition that could explain the late language onset. In fact, late talkers frequently present an average receptive vocabulary (comprehension vocabulary), and consequently, a larger expression-comprehension gap compared to same-age typical talkers. This latter group, without documented primary conditions, are the focus of this thesis, and so I will be referring to this group when I use the term 'late talker'.

Although most children identified as late talkers resolve during their first years of life, some language abilities remain weaker than same-age-children through school age to adolescence (Bishop & Edmundson, 1987; Bishop & Adams, 1990, Stothard et al., 1998; Rescorla, Roberts, & Dahlsgaard, 1997; Rescorla, Dahlsgaard, & Roberts, 2000; Rescorla & Roberts, 2002; Rescorla, 2002; Rescorla, 2005; Rescorla, 2009). Crucially, most of these skills are key for academic success, such as reading skills (Rescorla, 2005) or comprehension of complex sentences (Rescorla & Turner, 2015). Therefore, efforts are made to reduce or prevent early language delay through early interventions with the expectation of diminishing future social and academic failure. As soon as parents detect an expressive language delay in their child, they are typically advised to 'wait and see', a very criticised approach by some experts (see Singleton 2018). However, even when parents react rapidly and seek professional help, preventive measures are desirable to avoid any early language delay in the first case.

1.6 Measuring Language Emergence Through Parental Reported Inventories

To estimate the level of expressive language of young children, practitioners and researchers commonly make use of vocabulary checklists with normative data. For

American English speakers, one of the most frequently used is the MacArthur-Bates Communicative Development Inventories, (CDI; Fenson et al., 2007). Adaptations and translations of this inventory have been made to other languages such as Spanish (Jackson-Maldonado et al., 2003), Italian (Caselli & Casadio, 1995), and British English (Alcock, Meints, & Rowland, 2020), amongst others. The large amount of communicative and cognitive changes that young children experience between 8 and 30 months of age motivated the creation of two different versions of the CDI. The Words and Gestures CDI was designed for 8-to-18-month-old children. It evaluates the use of symbolic gestures and signs of communication, as well as the receptive and expressive vocabulary that typically emerges at this age. For children aged 16-30 months, the authors designed the Words and Sentences version of the CDI, which focuses on expressive language only, assessing vocabulary and first stages of grammar development.

It is common that researchers utilise a parental checklist to assign children with a vocabulary percentile according to their age and number of words produced. Late talkers are identified when they fall below or at a certain percentile, which, depending on the study, ranges between 5% to 30% (e.g., 5th percentile: Girolametto, Pearce, & Weizman, 1996; 10th percentile: Ellis Weismer et al., 2011; 15th percentile: Rescorla & Achenbach, 2002; 20th percentile: Beckage et al., 2011; 30th percentile: Jones, 2003). This is not always the case, though. Some researches use a criterion in which the child has to be older than a certain age to be considered a late talker (e.g., two years old in Rescorla, 1989; 18 months old in Fernald & Marchman, 2012). The rationale of the age cut-off resides in the high variability in vocabulary sizes at early ages and the fact that many late talkers catch up early on during infancy (Bates et al., 1994). After all, there is no diagnostic term such as "late talker" in any clinical manual, which makes it difficult to restrict the definition. In an attempt to avoid including those children at risk for DLD, some researchers included in their criteria a percentile cut-off for the receptive vocabulary. This is reasonable since late comprehenders have poorer vocabulary growth prospects compared to late talkers with better skills at word recognition (Fernald & Marchman, 2012). Nevertheless, the use of parental checklists is not only used for classifying children by their level of language attainment. The CDIs have the characteristic advantage of identifying commonality as well as individual differences which facilitates the investigation of processes involved in early language acquisition (Dale & Goodman, 2005).

One of the most salient advantages of parental checklists is that they are designed to be completed by parents with little assistance. This feature allows researchers to collect vocabulary data in a higher quantity and in a relatively short time. In general, parental reports have shown satisfactory reliability and validity (Fenson et al., 2000; Law & Roy, 2008). Yet, parental checklists are less accurate than in-lab methods for measuring vocabulary, and in many instances, accuracy varies depending on the parents' profile. For example, vocabulary checklists that were completed by more than one parent are more accurate (De Houwer, Bornstein, & Leach, 2005). Parents with a higher level of education are more accurate in reporting their children's language skills than parents with lower education levels, who tend to overestimate (Feldman et al., 2005). Parents with lower income are also likely to overestimate their children's language abilities (Fenson et al., 1993; Feldman et al., 2000). The report of productive vocabulary has shown high correlations with standardised behavioural language tests (for a review on studies that evaluated the concurrent validity of CDI see also Law & Roy, 2008). However, parents' abilities at estimating what their children can understand were suggested to be generally poor (Tomasello & Mervis, 1994; Moore et al., 2019), which challenges the validity of any study whose findings are supported with comprehension data obtained with CDI. Nevertheless, Alcock, Meints, and Rowland (2020) found that the new UK CDI Words and Gestures has excellent internal reliability for comprehension data (Cronbach's $\alpha = .99$) as well as good concurrent validity (comprehension data correlated significantly with results on object selection tasks: $r = .414$, $n=32$, $p = .019$).

Given the usefulness of CDIs in exploring the underlying mechanisms of early language acquisition as well as for the identification of late-talking toddlers, the CDI was chosen as the main instrument used to collect vocabulary data presented in the present thesis. In Chapter 4, the CDI was adapted to British English. Late talkers were identified by using a percentile threshold: 10th percentile in Chapter 2, and 20th percentile in Chapters 3 and 4. A higher cut-off was used for the latter chapters as it was desirable to assess the robustness of the results of another study that utilised the same type of analysis, and whose cut-off was 20th percentile (Beckage et al., 2011).

1.7 Language Delay and Children's Abilities

Various researchers have suggested that children's language skills are inheritable (e.g. Bishop, North, & Donlan, 1995; Rice, 2013). Whether linguistic abilities are

inherited or not, it is critical to understand how they develop over time in different circumstances to explain the diverse language trajectories observed in young children. Most studies that investigated the underlying causes of language delay in late talkers have focused on the abilities of children. The range of cognitive processes covered is extensive, and some of the studies concluded that there are small or no differences between late talkers and typical talkers. For instance, late-talking toddlers show similar amounts of joint attention bids (respond to or initiate joint attention) to younger typical talkers in structured situations where nonverbal communication is prompted (Vuksanovic & Bjekic, 2013). In a follow-up study, children identified as late talkers at age 2 obtained comparable average scores to typical talkers in attention, impulsiveness, and visuomotor planning at ages 4 to 6 (O'odorico & Assanelli, 2007). The ability to utilise cross-situational information for word learning seems to be unimpaired in late talkers as they benefit from a treatment based on this learning strategy (Alt et al., 2020).

In contrast, there is evidence from eye-tracking studies that suggested that late talkers might process novel words differently (Ellis et al., 2015). Late talkers initiate fewer joint attention bids than age-matched peers (Paul & Shiffer,1991); and the amount of attention bids predicts expressive and receptive language in typical talkers, but not in late talkers (Vuksanovic & Bjekic, 2013). Further, late talkers allocated shorter periods of attention to the to-be-learned objects (MacRoy-Higgins & Montemarano, 2016). Conversely, recent studies have made a strong connection between attention during word learning and the child's degree of shyness, with shyer children showing worse performance than typical talkers due to reduced attention towards the target object (Hilton, Twomey, & Westemann, 2019).

As well as atypical joint attention processes, late talkers were reported to exhibit shorter auditory sensory memory, which might impede the correct formation of phonological representations, and consequently influences word learning (Grossheinrich et al., 2010). In fact, phonological neighbourhood density in the vocabularies of late talkers is higher than same-age typical talkers (Stokes, 2010). That is, words in the late talkers' lexicon share more phonemes amongst them than the words in the typical talkers' lexicon, which might be related to a weaker ability to form new phonological representations. A weaker word learner has been suggested to exploit phonological density to facilitate language perception and production because words from dense phonological neighbourhoods tend to increase their mental activation as

compared to words that come from sparse phonological neighbourhoods (Scarborough, 2004).

In addition to weaknesses in those cognitive processes necessary for effective word acquisition, late talkers have been found to use well-known learning strategies in an atypical way. One of the most notable learning mechanisms is the *shape bias*, a strategy in which children extend known names for objects to novel instances with the same shape. Jones (2003), found that twelve 2-to-3-year-old late talkers showed no shape bias in a novel object name extension task, and showed a texture bias instead. Subsequent work has identified differences in the structure of late talkers' vocabularies that reflect the low shape bias usage, influencing the subsequent words they will learn (Perry & Kucker, 2019; Colunga & Sims, 2017). Beckage et al. (2011) suggested that late talkers might have a preference to learn words that have lower semantic relatedness with other words in the language environment. In their study, network analysis was used to characterise the vocabularies of 66 typical and late-talking children. Semantic relatedness of words, computed from word co-occurrence derived from the Child Language Data Exchange System database (CHILDES; MacWhinney, 2000), was used to connect the words in the child's vocabulary. Results showed that both typical and late talkers exhibit small-world structure, although late talkers present this to a lesser degree. The study suggests the existence of a relation between the child's rate of lexical development and the connectivity of her individual network. This finding led the authors to hypothesise the possibility of an 'oddball' strategy used by late talkers.

1.8 The Importance of the Characteristics of the Comparison Group

Most of the research that reported differences between late talkers and typical talkers shed some light on the potential causes of early language delay. For this purpose, it seems necessary to conduct comparisons of late talkers with age-matched typical talkers (e.g., from the research mentioned above: Ellis et al., 2015; Paul & Shiffer, 1991; MacRoy-Higgins & Montemarano, 2016; Grossheinrich et al., 2010; Stokes, 2010; Jones, 2003). However, when the question to answer is to what extent late talkers are simply 'slower' versions of typical talkers, vocabulary-size matched children are appropriate as a comparison group. To reject the slow version hypothesis, we need to eliminate the possibility of late talkers having general maturational delays that might be behind late

language onset. Would the studies just discussed above have obtained the same findings if the comparison group were vocabulary-matched instead? For instance, Colunga and Sims (2017) predicted that late talkers should show shape bias at ages 18 to 28 months. As the authors pointed out in their article, these children were nearly a year younger than the children tested in Jones' study (2003), suggesting that there is a chance that the participants in Jones' have outgrown the period in which shape is an important feature for word over-generalisation.

Matching children by vocabulary size in studies investigating language delay is useful for resolving the discussion on *categorical* versus *dimensional* accounts for language delay. Within this discussion, a first group of researchers argue that language learning mechanisms, potentially influenced by genetics (Bishop, 2006; Zubrick, Taylor, & Rice, 2012), might be impaired in some children, such as children with DLD (Rice et al., 2003), producing distinct language developmental outcomes that deviate from normative children (i.e., categorical account). As opposed to this support for qualitative differences between children with and without language delay, a second group of researchers defend the idea that quantitative differences can explain the individual variation of rates in language development (i.e., dimensional account). That is, for any language skill, children exhibit a specific level of endowment within a spectrum, with children with DLD being at the bottom of the ability continuum, which is normally distributed (Fernald & Marchman, 2012). With regards to late talkers, Rescorla and colleagues have provided consistent findings supporting the dimensional account through a collection of longitudinal studies (Rescorla, Roberts, & Dahlsgaard, 1997; Rescorla, Dahlsgaard, Roberts, 2000; Rescorla & Roberts, 2002; Rescorla, 2002; Rescorla, 2005; Rescorla, 2009). In these studies, participants were identified as late talkers at the age of two and were followed up until the age of 17. Many language-related skills were evaluated throughout these studies, and overall, children who were delayed at early years continued showing weaknesses in these skills. Results suggested that, although the language abilities typically develop as children mature, late talkers' abilities remain at the low end of the ability distribution according to their age.

Work like that of Rescorla and colleagues were important in providing critical evidence for the dimensional account of language development. Their findings have the theoretical implication that all children possess the same learning mechanisms that allow them to acquire language normally. This implies that the language skills of a child at the bottom of the spectrum according to her age is also at the average level of that language

skill for younger children. At the same time, this assumption also implies that the vocabulary profile of verbal-matched children should be identical or very similar. However, the existing literature that approached the dimensional question with verbal-matched controls provided mixed evidence for and against the dimensional account. In all the studies presented in this thesis that examine group differences, I explore the 'slow' versus 'deviance' hypothesis by matching children with language delay with younger children with the same level of language abilities (i.e., same vocabulary size).

1.9 Ability to Exploit Contextual Diversity in the Learning Environment

I have described earlier how contextual diversity can assist the child to form the mental word's concept by associating words with each other in a semantic manner. To do this, as well as general cognitive abilities, the child needs to be aware of the words occurring with other words within the speech stream. This ability relates to statistical learning, in which the child is sensitive to the distributional properties of words in the language environment. Not only is the child attentive to word-referent (mapping) occurrences but also word-word (or object-object) occurrences (for a detailed description of the components involved in statistical learning see Kachergis, Yu, and Shiffrin, 2009). With regards to late talkers, findings reported by Alt (2020) suggest that cross-situational learning is intact in late talkers; however, results from Beckage et al.'s (2011) suggest the opposite. In that study, the lower connectivity of the late talkers' networks might indicate an inefficient statistical learning ability to exploit contextual diversity cues in language. A potential explanation might be that word-word occurrences are harder to learn than word-referent occurrences. Unfortunately, not much literature has directly concentrated on understanding how children exploit contextual diversity for word learning, and therefore more studies are needed to fill this gap in the literature.

1.10 Language Delay and the Learning Environment

Parental education and socio-economic status (SES) are well-known predictors of later language outcomes (e.g., Reilly et al., 2007; Reilly et al., 2010; Rescorla, 2011; Fisher, 2017). An interesting study by Rowe (2008) found that the relation between SES

and the quality of child-directed speech is mediated by the amount of knowledge that caregivers have about child development. This finding implies that parents that are aware of the importance of a healthy upbringing tend to interact with their children in a way that favours their children's language outcome. If this is true, it is of importance to identify the parental behaviours that seem critical for early language acquisition. In particular, child-directed speech has been repeatedly reported as an important source of information for language learning (e.g., acoustic features: Dominey & Dodane, 2003; diminutive derivations: Kempe et al., 2001; repetitiveness: Newman, Rowe & Ratner, 2016; frequent frames: Mintz, 2003). Therefore, many efforts are dedicated to understand better how differences in quantity and quality of parental speech are associated with children's language rates as well as their language trajectories.

A prominent longitudinal study on young American children by Hart and Risley (1995) critically revealed the gap between children with low and high rates of vocabulary growth in the amount of linguistic input they received at home. Other studies also confirmed the link between the quantity of speech and later language achievements (e.g., Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Hoff & Naigles, 2002). An outstanding question is whether qualitative features of child-directed speech also play a role in facilitating language acquisition. Particularly, it is of special interest those aspects that are decisive in predicting future language delay. In a longitudinal study, Rowe (2013) investigated some maternal speech characteristics after controlling for SES, quantity of speech and early vocabulary skills of the children, and found that the number of word types and number of rare words are good predictors of later language outcomes. Newman, Rowe and Ratner (2016) reported that the repetitiveness (type-token ratio) and vocabulary diversity (word types) in maternal speech in nonverbal infants predicted the children's vocabulary at the age of 2. Thomas and Knowland (2014) modelled the acquisition of English past tense using a population modelling approach. External factors were modelled as the number of past tenses, types of regular verbs and types of irregular verbs in the speech. A reduction of these features was associated with the outcome ability levels of children with resolving delay.

A common linguistic variable that all these latter studies have underscored as important for predicting late language skills is the diversity of words in parental speech. Although these are important conclusions, how word diversity exactly facilitates language acquisition is less clear. A potential explanation seems to be in the relation between word types and contextual diversity. The larger the number of word types in

language, the more opportunities words have to co-occur with other word types. This consequently increases the contextual diversity degree of the words used in parental speech, which becomes then semantically richer. However, the diversity in parental speech increases as the child grows, as suggested by previous work (Hills, 2013). This is consistent with the idea that contextual consistency is more beneficial in early infancy for word learning (Roy et al., 2015). Nevertheless, it is relevant to investigate the natural ability of parents to gradually switch from consistency to diversity since keeping speech too simple might affect the child's opportunities to learn the different semantic features of words, an important part of word learning.

1.11 The Investigation of External and Internal Factors

The sections above have described the many different internal factors that have been investigated in order to shed light on the potential atypical underlying processes involved in word learning in late talkers. However, some researchers proposed that late talkers exhibit typical learning mechanisms, although weaker ones (i.e., they are slow learners). In order to contribute to the unresolved categorical versus dimensional debate, the present thesis investigates the semantic and syntactic composition of children's lexicons. As argued earlier, the use of vocabulary-matched comparison groups is suitable to answer the slow-versus-deviance question. Taking this into consideration, the children in Chapter 2 were grouped by vocabulary size to test whether the dimensional account was sustained when their lexical profiles were contrasted. This is the first study (to our knowledge) that directly compares the lexical profile of typical talkers, late talkers and children with autism spectrum disorder (ASD). Children with ASD represent an interesting population to investigate since most young children with ASD present early language delay (Charman, Drew, Baird, & Baird, 2003; Ellis Weismer et al., 2011; Mitchell et al., 2006). Similar to late talkers, the typical mechanisms for language learning have been reported to be intact in children with ASD (e.g., Shulman & Guberman, 2007; Naigles, Kelty, Jaffery, & Fein, 2011; Horvath, McDermott, Reilly & Arunachalam, 2018). Chapter 2 enquires whether typical ASD characteristics (e.g., social disinterest) interfere with normal language development by examining the type of words they learn in comparison to children without ASD. These are contrasted with the lexical profile of neuro-typical late talkers, to explore any commonalities in children

with delayed language onset, and to investigate their potential associations with age/developmental differences.

It has also been emphasised the potential impact that the child's language environment can have on early language development. No studies to date have investigated the impact of the quality of contextual diversity on language development. As explained earlier, the semantic structure of parental speech is an important source of semantic information that children exploit. Chapter 4 enquires whether differences in the semantic richness of child-directed speech (operationalised as contextual diversity) is associated with early language onset.

Results in Chapters 2 and 3 are derived from data that was obtained from a single point in time. Taking advantage of the large number of data points, we built cross-sectional developmental trajectories of the groups' vocabularies. This methodological arrangement allows a better understanding of the origin of the different developmental trajectories within the language-delay group (Thomas et al., 2009). Cross-sectional vocabulary trajectories are useful for hypothesis exploration and the production of preliminary results; however, longitudinal follow-up data are crucial in order to validate the types of changes which occur over time that were presumed previously. Consequently, the longitudinal study presented in Chapter 4 was conducted to further confirm conclusions made from the cross-sectional data in Chapter 3.

1.12 Aims and Outline of the Studies

In the remaining four chapters of this thesis, I present four independent studies that I carried out during my PhD. Although each chapter outlines their own individual research questions, each individual answer directly contributes to the general aim of this thesis: to characterise the internal and external factors involved in early word acquisition in relation to learning lexical meaning, and how these relate to the structure of the vocabularies of children with different rates of language development.

Chapter 2 presents a large-scale comparison of the early lexical profiles of children with ASD with that of typical and late-talking toddlers. The chapter aims to contribute to the dimensional-categorical debate by disentangling the differences in the lexical composition that are related to language delay from those related to the ASD characteristics. The outputs of this study contribute to our general aim in providing evidence of how differences in the lexical structure might be a reflection of internal

factors, particularly developmental changes associated with age and some unique ASD characteristics.

Chapter 3 examines the influence of semantic maturation in early lexical development. I investigated this influence with two studies. Study 1 compares the vocabularies of a large sample of typical and late talkers. The quantification of the network properties of the children's vocabularies indirectly reflects the degree to which children exploit contextual diversity cues in the language environment. Study 2 presents three computational models to investigate the potential role of semantic maturation in word development. For this purpose, I examine how the words' meaning is formed during the preverbal-comprehension phase, something that has been insufficiently investigated in the relevant literature. In the models, the simulated children leverage contextual diversity in the language learning environment to learn new words. The outputs of this study contribute to our general aim by confirming links between the internal cognitive components involved in word-meaning learning and the semantic structures of those vocabularies associated with different rates of vocabulary growth.

Chapter 4 presents a longitudinal study on the development of the semantic vocabulary structure of typical talkers, late talkers and late bloomers, as well as the inspection of the semantic structure in their respective language environments. Unlike the studies I outlined above which use secondary data, the study in Chapter 4 shows the result of a lengthy work I went through in designing the data collection instruments, promoting and recruiting families, collecting different types of longitudinal data remotely, and getting the audio files transcribed, all these throughout three years of my PhD. Importantly, the longitudinality of the vocabulary data serves as a validation for the findings derived from the studies in Chapter 3. The outputs of this study mainly contribute to our general aim by providing evidence for the latent associations between semantic quality in the linguistic input with the semantic structure of children's vocabularies, as well as with early language delay.

Chapter 5 investigates the role of contextual diversity and its counterpart, contextual distinctiveness in early word acquisition. These contextual metrics show opposite features of the word: how diverse it is across contexts, and how consistent are the contexts in which it appears. Since both contextual measures are known to predict word acquisition, I sought to shed some light on this paradox by conducting hierarchical regression models where the order of word acquisition is predicted.

Importantly, previous studies have differed in the way they measured contextual diversity to investigate its influence on word acquisition. I tested the predictive power of words' contextual diversity and contextual distinctiveness when measured in two different ways, and associated these with potential types of learning (semantic enrichment or word-referent mapping).

Finally, in Chapter 6 I summarise and discuss all findings described in Chapters 2, 3, 4 and 5 with the objective of providing a conclusive answer to the main intended question of the present thesis, which I pronounced above. I also discuss the theoretical and practical implications of the findings, followed by potential future directions and closing remarks.

Chapter 2: Identifying Areas of Overlap and Distinction in Early Lexical Profiles of Children with Autism Spectrum Disorder, Late Talkers, and Typical Talkers

This study compares the lexical composition of 118 children with autism spectrum disorder (ASD) aged 12 to 84 months with 4,626 vocabulary-matched typically developing toddlers with and without language delay, aged 8 to 30 months. Children with ASD and late talkers showed a weaker noun bias. Additionally, differences were identified in the proportion of nouns and verbs, and in the semantic categories of animals, toys, household items and vehicles. Most differences appear to reflect the extent of the age differences between the groups. However, children with ASD produced fewer high-social verbs than neurotypical children, a difference that might be associated with ASD features. In sum, our findings identified areas of overlap and distinction across the developing lexical profiles.

2.1 Introduction

Children with autism spectrum disorder (ASD) have significant delays in early language acquisition (Charman, Drew, Baird, & Baird, 2003; Ellis Weismer et al., 2011; Mitchell et al., 2006), but unlike neurotypical late talking children, these language delays are accompanied by restricted interests, repetitive behaviours and a social communication deficit (American Psychiatric Association, 2013). Might the language delay and the core deficits of ASD be related? This question highlights one of the central theoretical controversies within the ASD literature. That is, are the language delays associated with ASD merely adjustments along a continuum of development, where differences are primarily quantitative and along a single dimension (*the dimensional account*)? Or are the delays associated with ASD the result of a categorical difference in the way children with ASD learn language, giving rise to distinct language profiles that

are not simply delayed versions of typical development (*the categorical account*)? Similarly, are the language profiles of children with ASD similar to neurotypical late talkers, or do they represent a unique profile onto themselves?

Although the current diagnostic criteria for ASD does not include lexical or grammatical language deficits (American Psychiatric Association, 2013), receptive and expressive language delays have been found to differentiate children who will and will not go on to receive a diagnosis of ASD at ages as young as 12 months (Lazenby et al., 2016). Given this, previous research has examined the relation between various language domains and the language deficits in children with ASD (for an excellent review, see Eigsti, De Marchena, Schuh, & Kelley, 2011). Though previous work has looked at early developmental patterns of the lexicon among children with ASD (Charman et al., 2003; Luyster, Lopez, & Lord, 2007; Rescorla & Safyer, 2013, Ellis Weismer et al., 2011), the evidence needed to resolve the dimensional versus categorical account has been insufficient. The current study aims to address this problem by conducting an in-depth examination of the lexical composition of a large sample of children with ASD and to directly compare this with a large sample of children with typical language development as well as late talkers. Before we go on to describe our approach, we first describe the research supporting the dimensional and categorical accounts, lexical development in children with ASD and late talkers, and finally the putative role of social information in lexical development among children with ASD.

2.1.1 The Dimensional and the Categorical Account of Language Development

In the dimensional account of language development (Gernsbacher et al., 2005; Rescorla, 2009), children are placed along a continuum of language abilities, ranging from those with the poorest language skills to those with advanced language skills. Hence, the differences between a late talker and a typical talker are framed as being only quantitative (i.e., differences in the number of words produced), not qualitative (i.e., the type of words they produce). This account also implies that when late talkers and typical talkers are matched by language abilities (i.e., same number of words) the composition of their lexicons should remain similar. In contrast to the dimensional account, the categorical perspective of language development suggests that groups with language impairments demonstrate defining features of language development that do not align

with characteristics of typical language development (Dollaghan, 2004). In order to provide evidence for the categorical account, the identification of qualitative differences in the lexical profiles is useful because it can indicate the existence of potential atypical learning mechanisms. In this way, confirmation of lexical differences serves as guidance for future investigations of cognitive processes, providing further insight into potential categorical differences.

To date, many studies have provided evidence suggesting that children with language delay and typically developing children show similarities in their patterns of language development (e.g., Ellis Weismer ,2017; Rescorla, 2009). The same has been proposed for children with ASD with regards to the proportion of syntactic and semantic classes (Charman et al., 2003; Luyster, Lopez, & Lord, 2007; Rescorla & Safyer, 2013; Ellis Weismer et al., 2011). For instance, Charman et al., (2003) compared the proportion of words produced within syntactic classes (nouns, predicates, and closed-class words) in 87 preschool children with ASD to the normative sample for the MacArthur-Bates Communicative Development Inventory (CDI, Fenson et al., 1993). Charman et al. observed that the representation of the three syntactic classes across different vocabulary groups in the children with ASD was analogous to the pattern expected in a typical population. The proportion of semantic categories was also inspected in their sample. Children with ASD were reported to produce fewer words of the categories of 'Sound Effects', 'Animals', and 'Toys'; however, none of these differences was greater than 20% different relative to the CDI normative sample. In a later study conducted by Luyster, Lopez and Lord (2007), the percentage of syntactic classes was similar to that of typically developing children, even after controlling for verbal and nonverbal mental age, confirming the descriptive findings of Charman et al. (2003).

Rescorla and Safyer (2013) investigated the syntactic and semantic composition of early vocabularies of children with ASD by employing a different vocabulary inventory, the Language Development Survey (LDS, Rescorla, 1989). In their research, 45 children with ASD and 273 typically developing children were arranged into two overlapping groups by their total vocabulary: 1 to 49 words produced, and 1 to 310 words produced. Children with ASD and typically developing children who produced between 1 and 49 words had similar lexicons, for both syntactic and semantic classes. When examining the lexicons of the children who produced between 1 and 310 words, differences were found in the number of words produced in semantic categories; however, the

differences appeared to be explained by the overall lower vocabulary skills in the children with ASD relative to the normative comparison sample. Across the quantitative and qualitative analyses that Rescorla and Safyer (2013) conducted, many similarities were observed between the children with ASD and typically developing children, which suggested that the sample of children with ASD demonstrated a significant delay instead of deviance in lexical development.

The significant delay in lexical development in children with ASD frequently challenges researchers when attempting to control for age differences when comparing children with ASD with neurotypical children. Although previous work has documented that adults typically adapt their language input to the child's language level (e.g., Dykstra et al., 2012; Hani, Gonzalez-Barrero, & Nadig, 2013; Paul & Elwood, 1991), it is probable that older children are exposed to a somewhat different range of words which reflects changes in their immediate environment (e.g., "potty" instead of "diaper"). For this reason, an alternative comparison group to children with ASD is neurotypical late-talking children, who are closer in age. Although the majority of late talkers make significant language gains during the first years of life, many of them will experience persistent difficulties with some specific language abilities, such as in understanding and producing complex sentences at age five (Rescorla & Turner, 2015) and in non-word repetition tasks (Conti-Ramsden, Botting, & Faragher, 2001). Predicting future outcomes and vocabulary structure in late talkers has been the subject of much investigation (for a review, see Hawa & Spanoudi, 2014). For instance, Beckage and colleagues found that the structure of late talkers' vocabularies has less semantic clustering and is less tightly connected than vocabulary-matched typical talkers (Beckage, Smith, & Hills, 2011). Further, the emergence of word-learning biases has been computationally modelled in typical and late talkers' vocabularies to confirm the difference in the lexical structure of these two groups, such as a difference in the reliance on the shape bias (Colunga & Sims, 2017).

With regards to lexical composition, the percentage of the different syntactic and semantic categories in late talkers' vocabularies has been found to be similar to vocabulary-matched children, with the exception of the percentage of nouns, which has been found to be lower (MacRoy-Higgins, Shafer, Fahey, & Kaden, 2016). Ellis Weismer et al. (2011) compared 40 toddlers with ASD and 40 late talkers, who were matched on expressive vocabulary. The authors found no differences between the two diagnostic groups across the 18 semantic categories on the CDI. Noun proportions

were not examined in the sample; therefore, the question of whether the early vocabulary of children with ASD shows similar proportions of nouns to their neurotypical late-talking peers remained unanswered.

To date, a few studies on lexical composition give some weak support for the categorical account. Recent research has focused on individual lexical items within young children with ASD. In a large-scale study (209 toddlers with ASD and 272 neurotypical toddlers), Bruckner et al. (2007) observed that 25 words in the CDI are more likely to be learned by children with ASD (i.e., had a large bias). Bruckner et al. suggested that ASD symptomatology, such as restricted object use, deficits in orienting to social cues, and social communication deficits, might be related to vocabulary differences between children with ASD and typically developing children. A more recent study by Lazenby et al. (2016) also showed that certain words on the CDI were statistically more or less frequent in the vocabularies of infants who later were diagnosed with ASD, compared to typically developing infants.

Despite the insubstantial evidence gathered to support the categorical view of language delay, findings that identify different learning biases in children with ASD warrant the continued examination of evidence for the dimensional or categorical account of language development (e.g., Field, Allen, & Lewis, 2016; Happé & Booth, 2008; Pierce, Conant, Hazin, Stoner, & Desmond, 2011). Additionally, previous results from research that solely focused on the acquisition of nouns and verbs motivate us to further examine these two syntactic categories. For example, many studies have focused on a special case of lexical composition: the noun bias (e.g., Gentner, 1982). The greater percentage of nouns in early vocabularies not only has been observed in typically developing toddlers, but also in 2- to 3-year-olds with ASD (Swensen, Kelley, Fein, & Naigles, 2007). The noun bias has been linked to the well-known 'naming explosion' or spurt (Nelson, 1973; Benedict, 1979; Rescorla, 1980; Goldfield & Reznick, 1990). Many late talkers exhibited a reduced spurt, which suggests a potential link between noun acquisition and language delay (Rescorla, Mirak, & Singh, 2000). Different degrees of noun bias can be found in different languages, with the structure of the language being more influential in defining the intensity of noun bias than the parents' linguistic input (Dhillon, 2010). However, to our knowledge, previous research has not examined the possibility of identifying different degrees of noun bias and its relation to language abilities and ASD characteristics. The examination of the strength of noun bias seems relevant since previous studies have documented a weak or absent shape bias in

children with ASD and late talkers, an important learning strategy for early noun learning (Jones, 2003; Tek, Jaffery, Fein, & Naigles, 2008). In the present study, we will revisit the noun bias in the early vocabularies of children with ASD and late talkers with the aim to examine the strength of noun bias in these populations.

Although nouns are often the only syntactic class investigated in word learning studies, verbs have recently become the subject of interest among some researchers. Early verb acquisition may have a more important role in the later acquisition of grammatical abilities than nouns (Hadley, Rispoli, & Hsu, 2016). Some studies have focused on the type of verbs children acquire, which were classified according to syntactic features (transitive, intransitive and ditransitive; Olswang et al., 1997; Horvath, Rescorla, & Arunachalam, 2019) and to semantic features (manner and result verbs, punctual and durative verbs, number of event participants associated with its referent; Horvath, Rescorla, & Arunachalam, 2018; Horvath, Rescorla, & Arunachalam, 2019). Late talkers who showed less change in mean length of utterance (MLU) during a nine-week period produced fewer intransitive and ditransitive verbs than late talkers that showed greater MLU change (Olswang et al., 1997). Further, late talkers produced fewer manner verbs than their age-matched typical peers (Horvath, Rescorla, & Arunachalam, 2019). Regarding children with ASD, the syntactic bootstrapping strategies used to learn novel verbs by this population follow typical patterns (Shulman & Guberman, 2007; Naigles, Kelty, Jaffery, & Fein, 2011; Horvath, McDermott, Reilly, & Arunachalam, 2018). To our knowledge, the only other type of verbs investigated in children with ASD have been those that reflect mental states, which are described in the next section.

2.1.2 Social Interest Deficit and Word Acquisition

Deficits in social orienting among young children with ASD have been widely reported, including aspects such as responding less to their names or making less eye-contact (Osterling, Dawson, & Munson, 2002). Additionally, Pierce and colleagues showed that 14-month-old infants with ASD attended to moving geometric shapes longer than to children performing actions (Pierce, Conant, Hazin, Stoner, & Desmond, 2011). Children with ASD also have been found to show a higher preference for verbal and non-verbal noise over clear adult speech (Klin, 1991; Ceponiene et al., 2003). Different theories have suggested that this social disinterest in individuals with ASD either as a consequence of their deficits in social cognition (Social Cognitive Theory) or as a cause of their deficits in social cognition due to the diminished exposure to social

situations (Theory of Social Motivation; for a discussion contrasting both theories see Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012).

Studies have examined the potential ways in which social communication deficits and difficulties in understanding the social world influence word learning in children with ASD. Difficulties with understanding social intentions have been found to negatively influence the acquisition of verbs and prepositions (Parish-Morris, 2011). The acquisition of mental state verbs has been assessed (e.g., think, know, pretend) and suggested to be linked to weaknesses in Theory of Mind (Tager-Flusberg, 1992). Tager-Flusberg examined language samples from children with ASD and children with Down syndrome and found that children with ASD produced fewer mental state verbs. Ziatas, Durkin, and Pratt (1998) found that older children with ASD had poorer comprehension of mental state verbs than verbal-mental-age-matched children with Asperger syndrome, and neurotypical children with and without language impairment.

Horvath, Rescorla, and Arunachalam (2018) designed a word feature where verbs were linked to the number of participants that are usually associated with them. Horvath and colleagues found that typically developing toddlers are more likely to produce verbs that can describe scenes that involve fewer event participants than those that label scenes with more participants. The authors argued that verbs with fewer participants are easier to learn because the syntax in which they are embedded are easier to process. This word feature might be related to the degree of 'socialness' that children can perceive or be attracted to. In the current study, we explored this idea of words carrying social meaning. Verbs not only imply the number of event participants, but also the type of social interactions; for example, "smile" might evoke in the listener the act of someone smiling at someone else, or "share" might evoke someone sharing an object as part of a social interaction. Horvath, Rescorla, and Arunachalam (2018) demonstrated that typically developing children have greater difficulties in learning verbs that are associated with several event participants, one metric of the degree of socialness of the word. As such, given that children with ASD have difficulties attending to social cues, we wonder whether they would demonstrate pronounced challenges with learning highly-social words, relative to children who do not have ASD.

2.1.3 Current Study

Our main aim in the present study is to contribute to the dimensional-categorical debate by disentangling the differences in the lexical composition that are related to language delay from those related to the ASD characteristics. We conducted a large-scale comparison of the early lexical profiles of children with ASD with that of typical and late-talking toddlers to answer the following research questions:

1. *Do children with ASD and late talkers (LTs) show a noun bias to a similar extent as typical talkers (TTs) do?* To answer this question, we compare the relative difference between the proportion of verbs and the proportion of nouns between the talker groups. We hypothesised that the LTs and children with ASD might demonstrate a weaker noun bias, given that previous findings reported that these children do not demonstrate a shape bias (Jones, 2003; Tek, Jaffery, Fein, & Naigles, 2008).
2. *Do children with ASD, TT toddlers, and LT toddlers differ in the proportion of syntactic categories within their expressive lexicons?* To test for differences, we grouped children by vocabulary size, as has been similarly done by Rescorla and Safyer (2013), and Charman et al. (2003), and deferred to a fairly conservative statistical test corrected for multiple tests using Bonferroni alpha corrections to determine significance (see *Analysis Plan* section for details). In the case where differences exist in the proportion of syntactic categories, we also asked whether these differences were age-related. We examine the make-up of the differences to determine whether these differences are a result of the extent of language delay (in relation to age). Additionally, we identified the words that can be potentially affected by normal developmental changes in early childhood and then drew potential links between these words and the categories/classes where the differences were found. We tentatively predicted that TT children might produce more nouns relative to the other groups because of a robust shape bias.
3. *Do children with ASD show differences in the proportion of semantic categories compared to vocabulary-matched neurotypical children with and without language delay?* We followed a similar approach to addressing our first research question and provide a more detailed analysis description in the *Analysis Plan* section. As in our syntactic analyses, we also asked whether semantic differences were associated with the age differences across our groups. We predicted that the majority of the semantic categories would be similar across the groups; however, if differences did appear,

they would likely align with those identified by Charman et al. (2003; i.e., sound effects, animals, toys).

4. *Do children with ASD produce verbs with fewer social features than neurotypical children?* We collected a set of social word norms to directly evaluate the potential influence of social context in driving early verb learning differences among children with ASD and children who do not have an ASD diagnosis. Our concentration on verbs allows us to extend recent work by Horvath, Rescorla, and Arunachalam (2018) on semantic (social) features that were initially used to examine early verb learning in typically developing toddlers. Unlike Horvath and colleagues' measure of the number of participants associated to events and their associated actions, our social features rating has the advantage of capturing the degree to which verbs represent sociably acceptable behaviours, such as 'love' or 'hug', and those less socially accepted behaviours, such as 'hit' or 'hate'. We hypothesised that children with ASD would be less likely to be reported to produce verbs that are highly social.

2.2 Method

2.2.1 Participants

We examined early expressive vocabularies of 118 children with ASD from word-level data collected using the CDI, obtained from the National Database for Autism Research in January of 2019 (NDAR; Payakachat, Tilford, & Ungar, 2016). A comparison group of 4,688 typically developing children with CDI data was downloaded from a public repository, Wordbank (Frank, Braginsky, Yurovsky & Marchman, 2017) in September of 2018. We compared our ASD sample against LTs and TTs (see Table 2.1 for participant characteristics). Late-talking children were identified as those who fell at the 10th percentile or below on the CDI norms. This threshold has been used previously in relevant studies on LTs (e.g., Ellis Weismer et al., 2011; D'Odorico & Assanelli, 2007; Moyle, Ellis Weismer, Evans, & Lindstrom, 2007; Rescorla, 2009). The maximum number of words produced by LTs in our sample was 250. Therefore, our sample of 118 children with ASD was selected following the criteria that none of them had more than 250 words in their productive vocabularies. Even after this restriction, the LT sample had lower vocabulary sizes than the TT group (LT: M = 43.1, TT: M = 72.7; W = 674068, p < .001, d = 0.35, U_1= 24; description of the effect size statistic U_1 can be found in the *Analysis Plan* section) and the ASD group (*M*

= 74.9; W = 35030, p < .001, d = 0.32, U_1= 23). However, ASD and TT children had similar expressive vocabulary sizes (W = 230755, p = .30). In our analyses, we addressed this difference in expressive vocabulary size by subgrouping children according to the total number of words produced. In addition to vocabulary size, the ASD, LT, and TT children differed in age, with the ASD group being the oldest group, followed by the LT group, and finally the TT group (Age in months, ASD: M = 38.1, LT: M = 21.8, TT: M = 17.0; ASD vs LT: W = 50419, p < .001, d = 1.22, U_1= 63; ASD vs TT: W = 468383, p < .001, d = 0.54, U_1= 35; TT vs LT: W = 1592095, p < .001, d = 0.66, U_1= 41).

Data in our ASD sample were collected by different projects. The child data and study data for the children with ASD who were included in the current study can be inspected in NDAR, by searching our NDAR study DOI 10.15154/1518553. All 118 children were diagnosed with ASD by employing either Autism Diagnostic Observation Schedule (n = 61; ADOS, Lord et al., 1999), ADOS-2 (n = 10; Lord et al., 2012), Childhood Autism Rating Scale (n = 14; CARS, Schopler, Reichler, & Renner, 1988), Diagnostic Statistical Manual (n = 33; APA, 2013), or Autism Diagnostic Interview-Revised (n = 1; ADI-R, Rutter, Le Couteur, & Lord, 2003 – though several children received the ADOS and the ADI-R). Visual Reception and Fine Motor subscales of the Mullen Scales of Early Learning (MSEL; Mullen, 1995) have been widely used to assess non-verbal IQ. Given the extent of developmental delays in many children with ASD, age-equivalent scores are frequently reported (Bishop, Farmer, & Thurm, 2015; Bishop, Guthrie, Coffing, & Lord, 2011; Clark, Barbaro, & Dissanayake, 2017). Out of our sample, 98 children had a score for the Fine Motor subscale, with an age equivalent average of 26.8 months, ranging between 13 and 68 months; and 97 children had a score for the Visual Reception subscale, with an age equivalent average of 27.1 months, ranging between 11 and 54 months. The age equivalent of the ASD group (M = 26.9, SD = 7.03) is still significantly higher than the chronological age of LT group (W = 34134, p < .001, d = 0.6, U_1= 38), and TT group (W = 372424, p < .001, d = 0.4, U_1= 27; age equivalent is an average of the two subscales; the 21 children without MSEL scores were excluded from this group comparisons).

Table 2.1 Participant Characteristics

Diagnostic Group	Number of children	Productive vocabulary size average, *range,* (*SD*)	Age average[1], *range,* (*SD*)	Checklist used
ASD	118	74.9 *1-250* (75.7)	38.1 *12-84* (15.9)	W&G = 31 (26.3%) W&S = 87 (73.7%)
TT	4142	72.7 *1-250* (66.3)	17.0 *8-29* (3.4)	W&G = 1739 (42%) W&S = 2403(58%)
LT	484	43.1 *1-248* (51.1)	21.8 *16-30* (4.6)	W&G = 28(5.8%) W&S = 456(94.2%)

Note. Typically Developing Children = TDC; Typical Talkers = TT; Late Talkers = LT; Autism Spectrum Disorder = ASD. CDI forms: W&G = Words and Gestures form, W&S = Words and Sentences form; [1]Age is in months.

2.2.2 Syntactic Classes and Semantic Categories

The vocabularies of the children in the present study were assessed using two versions of the CDI: the CDI – Words and Gestures, normed on children between 8 and 18 months, and the CDI - Words and Sentences, normed on children between 16 and 30 months. To compare the early vocabularies of our groups of children, we conducted separate analyses according to semantic categories and syntactic classes.

For the syntactic analysis, we examine the two main types of words: nouns and verbs. Our motivations to study these two syntactic groups are based on the relevance that previous research established between their early acquisition and later language abilities (Benedict, 1979; Hadley, Rispoli & Hsu, 2016; Goldfield & Reznick, 1990; Nelson, 1973; Rescorla, 1980; Rescorla, Mirak, & Singh, 2000). Nouns consisted of the words that were contained in the following CDI categories: Animals (43 words), Vehicles (14), Toys (18), Food and Drink (68), Clothing (28), Body Parts (27), Furniture and Rooms (33), and Small Household Items (50). As in Bates et al.'s (1994) classification for the syntactic class of nouns, we excluded the following categories because it has been suggested they do not follow the typical growth of 'true nominals' (Snyder, Bates, & Bretherton, 1981; Bates et al., 1994): Sound Effects and Animal

Sounds (12), Outside Things (31), Places to go (22), People (29), Games and Routines (25). The verb class included words classified as Action Words (103).

For the semantic analysis, all the CDI categories were considered. These are, in addition to the CDI categories mentioned so far: Descriptive Words (63), Pronouns (25), Questions words (7), Prepositions and Locations (24), Quantifiers and Articles (17), Words about Time (12), Connecting words (6) and Helping Verbs (21). The items "inside/in" from the CDI Words and Sentences and "in" and "inside" from the CDI Words and Gestures were not used because when the item "inside/in" was marked it was not clear enough to confirm whether the child said one or two words. We decided to analyse those semantic categories excluded in our syntactic analysis for two reasons. Firstly, a considerable proportion of words in early vocabularies are composed of words from these semantic categories. Although these words have been suggested not to follow a typical nominal growth, we believe that any word from these categories could potentially be subject to an age effect, a factor of interest in the current study. Secondly, since we are building on previous studies that examined all the CDI categories (Charman et al., 2003; Luyster et al., 2007), we sought to analyse the same categories, which were comprised of the same word types, to contrast our results with those of these studies.

For each child, we computed the proportion of words produced for each syntactic class and each semantic category given the child's total expressive vocabulary size. To calculate the expressive vocabulary size, we considered all words reported to be produced on the CDI. Then we sub-grouped the samples into bins according to the total number of CDI words produced. This approach allowed us to examine whether different patterns arise across different points of vocabulary development. This approach was particularly important because Bates and colleagues (1994) have suggested that vocabulary sizes between 1 and 25 or even 50 words are unstable. Furthermore, subgrouping the samples into bins enabled us to control for differences in vocabulary size in our LT group. Therefore, for up to 100 words, we use bins covering a range of 25 words. Between 101 and 250 words, we use bins covering a range of 50 words.

2.2.3 Words Influenced by Developmental Stages in Early Childhood

To identify those words that are potentially associated to specific developmental stages throughout early childhood, we first split the word-level data into two age groups: the TTs as the 'younger' group, and the LTs and children with ASD as the 'older' group. Then, for each vocabulary bin, we computed the proportion of younger children and older children that produced each item/word separately for each group. Next, we subtracted the word proportions of the younger children from that of the older children per vocabulary bin. These subtractions resulted in positive numbers, which identify those words that younger children produced more often than older children, and negative numbers, which identify those words that older children produced more often than younger children. Finally, we extracted the top-10 most negative words and top-10 most positive words for each of the eight vocabulary bin comparisons.

We conducted a post-hoc examination of the words identified following the procedure just described, focusing only on those that belong to the categories where differences between LTs, TTs and children with ASD were previously found in our semantic analysis. Our objective was to determine whether the proportion of these semantic categories were related or not to developmental differences between the groups. To do this, we examined some word features associated with developmental changes that occur in early childhood. We concentrated on physical development as these changes are likely to influence the presence or absence of objects in the child's environment, as well as the relation that children have with the objects. We expect to identify words like "diaper" and "bottle" in younger children, and words like "potty" and "fork" in older children. In the case of verbs, it is particularly challenging to infer how changes in physical development can influence verb acquisition as young children are able to learn the meaning of actions by observing other people (Huttenlocher, Smiley, & Charney 1983). In addition, verbs can be learned as events and not actions, for example, "walk" can be understood as the event of going to the park instead of the act of walking. However, we focused on the social aspect embedded in verbs, which we describe in the following section.

2.2.4 Social Features in Verbs

In order to examine whether features associated with the core deficits of children with ASD influence early vocabulary development, we examined the social features of the words listed under the CDI Words and Sentences Action Words category. Social ratings for each verb were collected from a sample of 54 adults using a survey that was distributed on Amazon's Mechanical Turk platform. The participants lived in the United States of America and self-reported to be native English speakers. Thirty-one participants identified as male, 21 identified as female, and two participants identified as "other". The average age of the participants was 35.9 years (range: 22 - 72 years), and the average reported household income was $47,444 (range: $7,000 - $120,000). The sample was 83.3% White, but also included four Asian individuals, two Black individuals, and one "other". Additionally, four individuals reported to be Hispanic.

The participants were given the following prompt "For each of the following words, please type in a number between 1 and 10 to rate how social each word is (1 = Not Social, 10 = Extremely Social). A word is more social if it typically involves interacting with other people. A word is less social if it typically does not involve other people." This approach to measuring social features of the verbs was similar to the approach used by Horvath, Rescorla, and Arunachalam (2018). The order of the words was pseudo-randomised so that they were not in alphabetical order. Three items were added to the survey to test for attention by asking the participant to select a specified word from a list of three words. Every participant passed these items; therefore, no participants were excluded. The average social rating for each verb was calculated. Then, we calculated the median social rating score for the verbs reported to be produced by each child in our sample. Following this, to control for the higher proportion of verbs that our sample of children with ASD produced, we subdivided our sample according to vocabulary bins of verbs produced: 1 to 25 verbs, and 26 to 50 verbs. Not all children in our sample were included in this analysis since some children with small vocabularies produced no verbs. The subsample analysed comprised of 83 children with ASD, 233 LTs, and 2,457 TTs.

2.2.5 Analysis Plan

We chose to conduct non-parametric pairwise comparisons using the Wilcoxon rank-sum test to test for group differences since the distribution of proportions across

vocabulary sizes violated the assumption of homogeneity of variance. To be able to control for vocabulary size, we tested the children within each vocabulary size bin (see Table 2.1 for vocabulary sizes analysed). The Benjamin and Hochberg false discovery rate procedure was implemented in each test performed. Additionally, all *p*-values were corrected using the Bonferroni method as we conducted several distinct analyses on the sample; corrected *p*-values are reported.

In order to explore the influence of age on word proportions, we considered any difference that emerged between the three groups in one or more vocabulary bins. Since many semantic categories are composed of a few words, there is a chance that differences between two groups emerge in some of them; therefore, to minimise type 1 error, we will only acknowledge differences between groups if we found significant differences in at least two vocabulary bins. To evaluate the effect size of the significant results, we report two statistics following the suggestion made by Fritz, Morris and Richler (2012). The first statistic is the well-known Cohen's *d*, which we interpret following the Cohen's convention (1988). In addition, to facilitate the interpretation of the relation between the groups' distributions, we also report U_1, also created by Cohen (1988), which is the percentage of non-overlap between the two distributions.

2.3 Results

2.3.1 Noun bias

The LT and ASD groups showed a higher proportion of nouns (ASD: *Mdn* = 0.47; LT: *Mdn* = 0.37) than verbs (ASD: *Mdn* = .08; LT: *Mdn* = 0.007; ASD: $Z = 8.4$, $p < .001$; LT: $Z = 18.52$, $p < .001$). Typical talkers also showed a higher proportion of nouns (*Mdn* = .48) than verbs (*Mdn* = .04; $Z = 8.21$, $p < .001$). The effect size of all noun bias analyses were very large (ASD: $d = 2.0$, $U_1 = 81$; LT: $d = 2.70$, $U_1 = 90$; TT: $d = 3.58$, $U_1 = 96$). To further explore this noun bias, we subtracted the proportion of verbs from the proportion of nouns for each child in our sample. Then, we compared these verb-noun gaps between our three groups in a subsample of the children with vocabulary sizes between 1 and 75, which includes the vocabulary bins where nouns and verb differences were found in our syntactic analysis (reported in the next section; ASD: 77 children; LT: 400 children, TT: 2,646 children). The ASD group and the LT group had the smallest verb-noun gaps (ASD: *Mdn* = .27; LT: *Mdn* = .30), and the TT

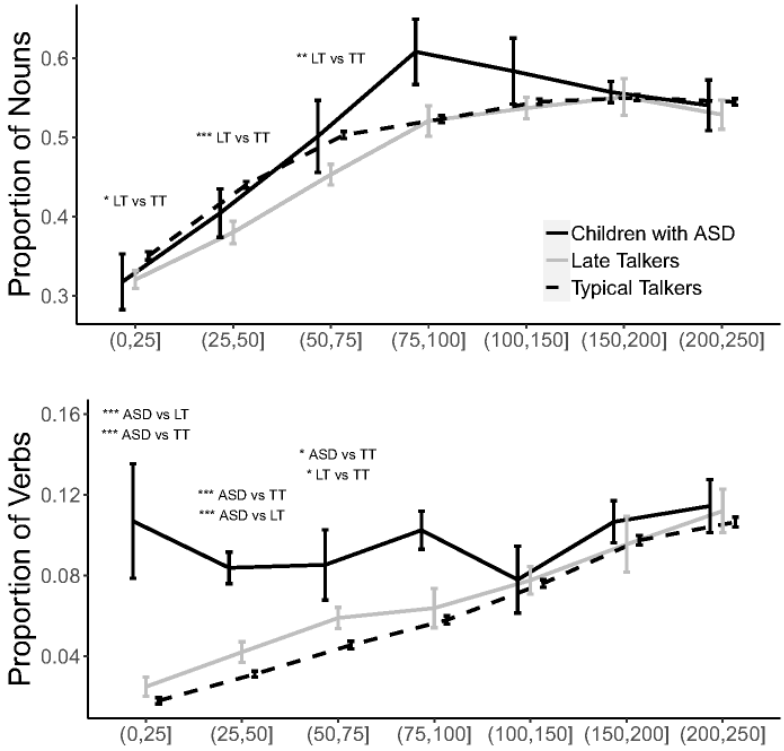
group showed the largest gap (*Mdn* = .39). The ASD group differed from the TT group (W = 76278, p < .001) but not from the LT group (W = 13995, p > .05). There were significant differences between the TT group and the LT group (W = 413471, p < .001). The effect sizes of the noun-verb gap analyses were generally small (ASD vs TT: d = 0.1, U_1= 8; LT vs TT: d = 0.26, U_1= 21).

2.3.2 Syntactic Classes and Semantic Categories

We first compared the early production of nouns and verbs across children with ASD, TTs, and LTs. Figure 2.1 depicts the proportion of words produced within each syntactic class according to vocabulary bin size for the ASD, TT, and LT groups. Our analysis on the proportion of nouns revealed that there were no differences between children with ASD and either LTs or TTs (See Table 2.2). Late talkers produced a lower proportion of nouns than TTs for vocabulary sizes 1 to 25, 26 to 50, and 51 to 75. The effect sizes of the noun analyses were generally small (d = [0.1,0.3]; U_1= [8, 21]). Verbs revealed a different pattern. Children with ASD produced a higher proportion of verbs than TTs at vocabulary bins 1 to 25, 26 to 50, and 51 to 75; and produced a higher proportion of verbs than LTs for the two smallest vocabulary size bins. Late talkers exhibited a larger proportion of verbs than TTs for the 51-75 vocabulary size bin. The effect sizes of the verb analyses range from small to medium for the ASD vs TT comparisons (d = [0.2,0.4]; U_1= [15, 27]), medium to large for the ASD vs LT comparisons (d = [0.4,0.9]; U_1= [27, 52]), and small for the TT vs LT comparison (d = 0.3; U_1= 21). Given that the Bates et al. (1994) syntactic classifications have been frequently used in previous research (e.g., Braginsky, Yurovsky, Marchman, & Frank, 2019; Charman et al., 2003; Luyster et al., 2007; MacRoy-Higgins, Shafer, Fahey, & Kaden, 2016; Roy, Frank, DeCamp, Miller, & Roy, 2015), we also classified the words following their approach. Bates et al. (1994) categorised CDI words as either nouns, predicates (adjectives and verbs), or closed-class words. Analyses of these syntactic classes replicated the significant differences between the groups for the nouns and verbs/predicates and revealed no differences for closed-class words (see Appendix A).

In addition to examining the syntactic organisation of early vocabulary development, we inspected the 22 semantic categories on the CDI. We found differences between the groups in the semantic categories of Action Words, Animals, Small Household items, Toys and Vehicles. We will not discuss Action Words since

Figure 2.1 Proportion of Syntactic Classes in the Vocabulary of Children with ASD, Late Talkers and Typical Talkers.



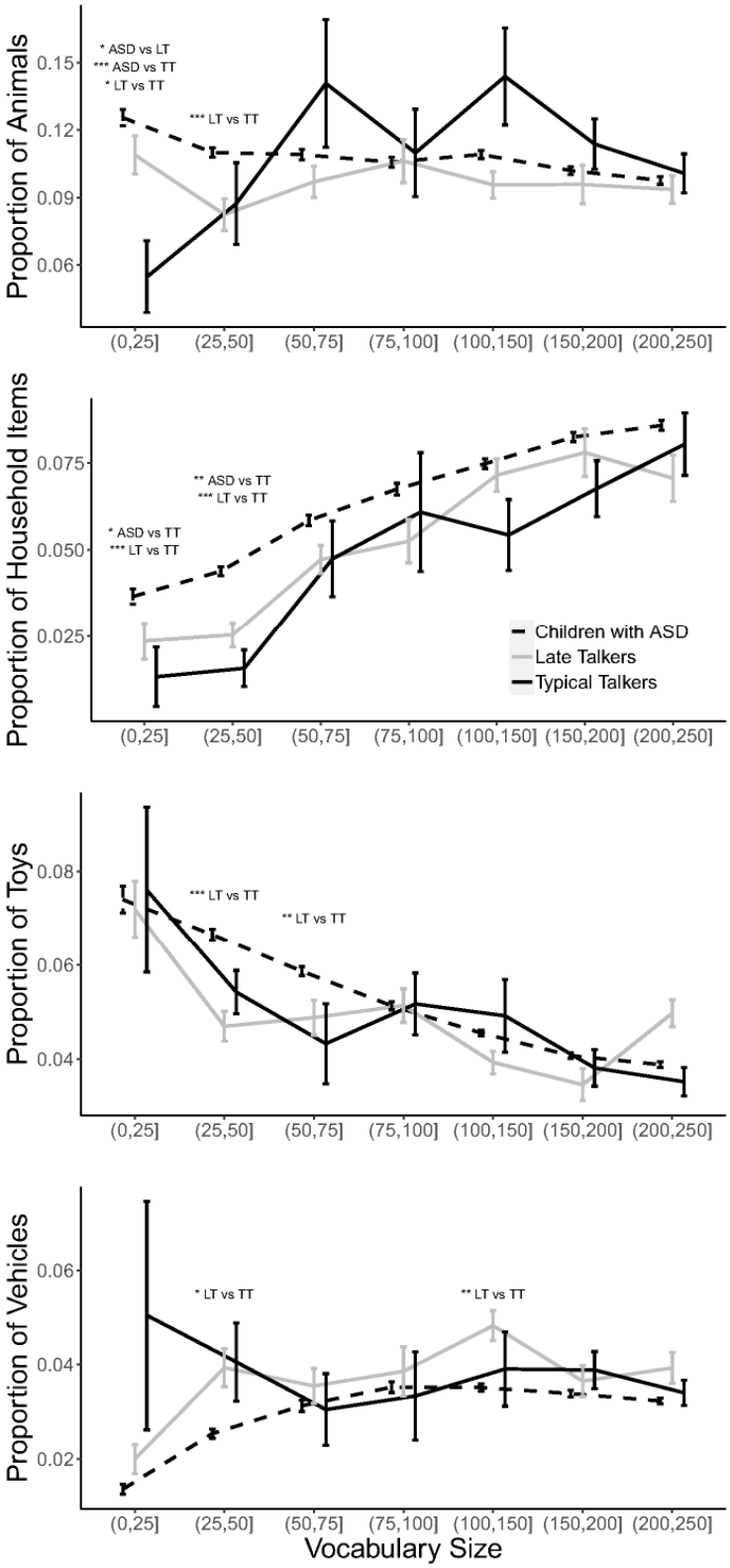
Note. Error bars, signifying standard error of the mean, have been shifted slightly to facilitate visibility. Asterisks indicates a significant group difference: * $p < .05$; ** $p < .01$; *** $p < .001$.

verbs were already discussed for the syntactic analysis. The following results are shown in Figure 2.2 and Table 2.3. We found significant differences in the proportion of animal words between all groups at vocabulary sizes 1 to 25, where TTs presented the highest proportions, followed by LTs and then the children with ASD with the lowest proportion. In the vocabulary bin of 26 to 50 words, LTs and children with ASD produced a similar proportion of animal words, and only LTs significantly differed from the TTs, with TTs producing the highest proportions. In our small household items findings, LTs and children with ASD presented similar proportions at small vocabularies only. Both differed from the TTs who produced more of these words in vocabulary sizes 1 to 25 and 26 to 50 words. Our toy words finding revealed that LTs and children with ASD had similar proportions of toy words, but only LTs differed from the TTs in vocabulary bins 26 to 50, and 51 to 75, with TTs showing the highest proportions. Similarly, LTs and children with ASD showed similar proportions of

vehicle words, with LTs being the only of the two groups to differ from the TTs, who produced more vehicle words at vocabulary sizes 26 to 50, and 101 to 150 words. Overall, out of the seven vocabulary sizes analysed, differences were found only in two vocabulary sizes in each semantic category that we discussed. These were mostly small vocabulary sizes and mostly between LTs and TTs. In addition, the differences found had a small effect size (ASD vs TT: d = [0.2, 0.3]; U_1 = [15, 21]; LT vs TT: d = [0.2, 0.3]; U_1 = [15, 21]), and one difference had medium effect (ASD vs LT: d = 0.4 ; U_1 = 27).

In contrast to our syntactic analysis, our semantic analysis presents the disadvantage that many of the CDI categories are composed of a small number of words. Although the statistical tests detected significant differences in proportions between the groups, it is necessary to consider the relative size of the differences in terms of the number of words that children usually produce at the corresponding vocabulary sizes. In our sample, 90% of children with a vocabulary size between 1 and 50 words produced between 0 and 2 small household items (1 to 25 words: M = 0.41, SD = 0.70; 26 to 50 words: M = 1.57, SD = 1.46), and between 0 and 5 animals (1 to 25 words: M = 1.51, SD = 1.52; 26 to 50 words: M = 3.96, SD = 2.41); 94% of the children with vocabulary sizes of 26 to 75 words produced between 0 and 4 toys (26 to 50 words: M = 2.35, SD = 1.27; 51 to 75 words: M = 3.54, SD = 1.36). Regarding vehicle words, 90% of children with vocabulary sizes of 26 to 150 words produced between 0 and 5 vehicles (26 to 50 words: M = 1.02, SD = 1.16; 101 to 150 words: M = 4.50, SD = 2.27).

Figure 2.2 Proportion of Animals, Small Household Items, Toys and Vehicles in The Vocabulary of Children With ASD, Late Talkers and Typical Talkers.



Note. Error bars, signifying standard error of the mean, have been shifted slightly to facilitate visibility. Asterisks indicates a significant group difference: * $p < .05$; ** $p < .01$; *** $p < .001$.

Table 2.2 Wilcoxon Rank Sum Test for Nouns and Verbs.

					ASD vs TT			ASD vs LT			TT vs LT		
Syntactic Class	Vocabulary Size	Mean ASD	Mean LT	Mean TT	*W*	*p*	*d*	*W*	*p*	*d*	*W*	*p*	*d*
Nouns	1-25 words	0.32	0.32	0.35	24914.5	>.05	-	5707.5	>.05	-	193148.5	<.05	0.1
	26-50 words	0.40	0.38	0.44	7179.0	>.05	-	834	>.05	-	42368	< 0.001	0.3
	51-75 words	0.50	0.45	0.50	3005.0	>.05	-	413.5	>.05	-	17023	< 0.01	0.3
Verbs	1-25 words	0.11	0.025	0.018	40113.0	< .001	0.2	8092.5	< .001	0.4	169001	>.05	-
	26-50 words	0.084	0.042	0.031	15090.0	< .001	0.4	1221.5	< .001	0.9	28761	>.05	-
	51-75 words	0.085	0.059	0.046	4167.5	< .05	0.2	423.5	>.05	-	10004	<.05	0.3

Note. Groups considered in the analysis are the ASD group, LT (Late talker) group, and TT (Typical talker) group. Only the three vocabulary bins that were significant are displayed. All *p*-values were first corrected using the Benjamini-Hochberg procedure (BH), then corrected again using the Bonferroni method (i.e., corrections accounted for comparisons for the three syntactic classes).

Table 2.3 Wilcoxon Rank Sum Test for Differences Across Semantic Categories

					ASD vs TT			ASD vs LT			TT vs LT		
Semantic Category	Vocabulary Size	Mean ASD	Mean LT	Mean TT	*W*	*p*	*d*	*W*	*p*	*d*	*W*	*p*	*d*
Animals	1-25 words	0.055	0.11	0.13	16696	< .001	0.3	4309.5	< .05	0.4	194996.5	< .05	0.2
	26-50 words	0.087	0.083	0.11	6392	> .05	-	784.5	> .05	-	43113	< .001	0.3
Small Household Items	1-25 words	0.013	0.024	0.037	22106.5	< .05	0.2	5697	> .05	-	204735.5	< .001	0.2
	26-50 words	0.016	0.025	0.044	4594.5	< .01	0.3	649	> .05	-	44014.5	< .001	0.3
Toys	26-50 words	0.054	0.046	0.067	6762	> .05	-	910	> .05	-	44596.5	< .001	0.3
	51-75 words	0.043	0.049	0.059	1997	> .05	-	322	> .05	-	17341	< .01	0.3
Vehicles	26-50 words	0.041	0.049	0.025	10798.5	> .05	-	781.5	> .05	-	25856	< .05	0.2
	101-150 words	0.039	0.048	0.035	3238.5	> .05	-	184.5	> .05	-	5333.5	< .01	0.3

Note. Results obtained from post-hoc Wilcoxon rank-sum tests for the four semantic categories that showed significant differences between the groups: Animals, Small Household Items, Toys and Vehicles. Groups compared in the analysis are the ASD group, LT (Late talker) group, and TT (Typical talker) group. Only the vocabulary bins that were significant are displayed. All *p*-values were firstly corrected using the BH method then corrected again using the Bonferroni method (correcting for comparisons across 22 semantic categories).

2.3.3 Words Associated with Developmental Stages

To examine the potential influence of age in our findings of lexical differences, we explored word-level differences. Table 2.4 shows the set of words whose production are potentially influenced by age in children at similar stages of vocabulary development. We only display and discuss those CDI categories where differences were identified in our semantic analysis; however, the full results can be found in Appendix B. The number of words that we identified reflect the proportions that the corresponding comparison groups showed in our semantic analysis; that is, more words were identified in the group that previously showed higher proportions in the respective CDI category.

We found age differences in words related to the physical readiness for potty training, and the development of the digestive system. Regarding small household items, two of the words that younger children are more likely to produce, "bottle" and "spoon", seem to be related to early stages of feeding. In our examination, we also identified other words that we expected to be affected by age, but that belong to other semantic categories where group differences were not observed. Some examples are "bib" or "cracker" in the younger group, and "go potty" and "candy" in the older group.

We also explored features related to fine versus gross motor skills. These features can be applied to those words that represent objects that children are normally allowed to manipulate, i.e., toys. In this sense, we expect that older children would engage more often in playing with toys where fine motor skills are required. In the case of the toy-related words, which are composed of the CDI categories of Animals, Toys, and Vehicles, we identified words that require fine motor skills in both groups (older group: "play-dough"; younger group: "block"). Interestingly, the older children are likely to produce toys with small rotating parts ("bus", "car", "helicopter", "motorcycle", "truck"), which contrast with the high proportion of toys produced by the younger group that are generally characterised for the lack of small mobile pieces ("doll", "teddy bear", "balloon").

Table 2.4 Words that Older and Younger Children Produce which are Potentially Related to Differences in Development.

Animals		Small household Items		Toys		Vehicles	
Younger children	Older children	Younger children	Older children	Younger children	Older children	Younger children	Older children
bear (*c*, *e*)	bee (*c*)	bottle (*a*, *b*, *c*, *d*)	plate (*f*)	balloon (*a*, *b*, *c*)	ball (*a*)	airplane (c)	bus (*e*)
bird (*a*, *b*, *d*)	chicken (*g*)	broom (*f*)		block (*e*)	playdough (*g*)		car (*a*, *b*)
bunny (*b*)	horse (*d*)	keys (*d*)		book (*a*, *b*)			helicopter (*e*)
cat (*b*)	lion (*g*)	spoon (*e*)		doll (*e*)			motorcycle (*e*)
dog (*b*)	penguin (*g*)			toy (*f*)			truck (*d*)
duck (*b*)							
kitty (*a*, *b*)							
owl (*g*)							
pig (*g*)							
teddy bear (*c*)							

Note. Letters in parenthesis represent the vocabulary size where the word was identified. *a*: 1 to 25 words; *b*: 26 to 50 words; *c*: 51 to 75 words; *d*: 76 to 100 words; *e*: 101 to 150 words; *f*: 151 to 200 words; *g*: 201 to 250 words.

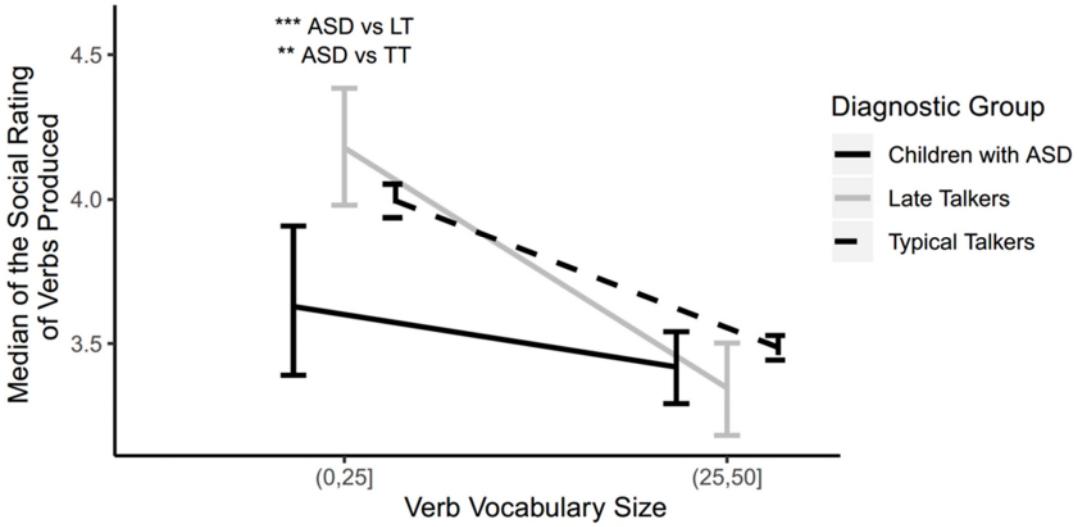
2.3.4 Social Features and Verb Acquisition

Regarding the social ratings for verbs given by adults, the highest rating value was 9.2 and corresponded to the words "kiss" and "hug"; other highly rated words included "help" (8.3) and "tickle" (8.1). The lowest social rating value was 1.6 and corresponded to the word "rip"; other words that received low social ratings included "sweep" (1.8) and "jump" (2.0). The average social rating value was 3.9; words with similar scores are "see" (3.7), "cry" (4.0), and "cook" (4.2). Low socially acceptable actions that involve more than one participant, like "hit" and "hate", received medium rating values (4.6 and 4.2 respectively). Apart from "kiss", "hug" and "help", other high socially desirable actions received high scorings, such as "share" (8.6) and "love" (8.5).

To address our last research question—whether characteristics associated with ASD symptomatology, particularly deficits in social abilities, relate to verb production—we examined the relationship between verb acquisition and the degree to which verbs are associated with social interactions. We compare children with the same number of verbs in their vocabularies. One group of children produced between 1 and 25 verbs (TT n = 2457, LT n = 233, ASD n = 83), and the other group of children produced between 26 and 50 verbs (TT n = 162, LT n = 9, ASD n = 9). As can be observed in Figure 2.3, children with ASD within the first 25 verb bin produce verbs with significantly lower social ratings (Mdn = 3.1) when compared to TTs (Mdn = 3.6, W = 80239, $p < .01$) and to LTs (Mdn = 3.7, W = 6931, $p < .001$). No differences were found between the LTs and TTs (W = 264746, $p > .05$). The effect size was small for the ASD vs TT comparison (d = 0.13; U_1 = 10), and medium for the ASD vs LT comparison (d = 0.44; U_1 = 30).

From the developmental point of view, Figure 2.3 suggests that all groups generally produced more high-social verbs at the early stages of

Figure 2.3 Median Social Rating Scores for Verbs that Children with ASD, Late Talkers and Typical Talkers Produced Relative to the Total Number of Verbs Produced.



Note. Error bars, signifying standard error of the mean, have been shifted slightly to facilitate visibility. Asterisks indicates a significant group difference: ** $p < .01$; *** $p < .001$.

vocabulary development (i.e., at vocabulary sizes 1 to 25 words). However, only TTs showed significant differences between the two verb vocabulary bins (verb vocabulary size (0,25] *Mdn* = 3.6, *M* = 4.0; verb vocabulary size (25,50] *Mdn* = 3.5, *M* = 3.5; $W = 224078$, $p < .01$). This difference had a small effect size ($d = 0.11$; $U_1 = 9$). LTs and children with ASD did not differ across the two vocabulary sizes analysed (LTs: verb vocabulary size (0,25] *Mdn* = 3.7, *M* = 4.2; verb vocabulary size (25,50] *Mdn* = 3.3, *M* = 3.5; $W = 1388$, $p = .09$; ASD: verb vocabulary size (0,25] *Mdn* = 3.1, *M* = 3.6; verb vocabulary size (25,50] *Mdn* = 3.4, *M* = 3.4; $W = 324.5$, $p > .05$).

2.4 Discussion

The current study identified group differences across syntactic classes, semantic categories, and social features of verbs. Our findings highlighted group differences that primarily occur at the earliest stages of lexical development. In what follows, we discuss our findings in the context of the broader literature.

2.4.1 Noun Bias

The children with ASD showed a noun bias in their vocabularies, supporting previous findings (Swensen, Kelley, Fein, & Naigles, 2007). However, in our attempt to measure the strength of the noun bias (proportion of verbs subtracted form proportion of nouns, or noun-verb gap) we found that LTs and children with ASD showed a weaker noun bias compared to TTs. In addition, LTs and children with ASD showed similar sizes of the noun-verb gap. Within the noun bias literature, it has been determined that the language spoken at home is the main factor that drives the strength of the noun bias (Dhillon, 2010). Since all the children in our sample were English speakers, we can discount this effect and posit that late language onset is a factor that regulates the size of the noun-verb gap. To help interpret the noun-verb gap findings, it is helpful to consider the syntactic comparisons. We found medium to large effect sizes observed in our verb comparisons. Although all of the groups demonstrated a noun bias, the comparatively high proportion of verbs that the children with ASD produce mainly influences the size of the noun-verb gap difference relative to the TTs. However, it is important to remember that the effect sizes of the noun-verb gap group differences were generally small; as such, these findings underscore the consistency of the noun bias across groups but identify interesting differences in its degree.

2.4.2 Syntactic Classes and Semantic Categories

Within our comparison of the proportion of the syntactic classes, the verb differences were the most striking. The ASD group, which consisted of older children with the largest language delays showed the highest proportions of verbs, followed by the second oldest group (LTs) and finally the youngest group (TTs) with the lowest proportions. A possible explanation might be that, since LTs and children with ASD are older, their cognitive abilities are more mature than verbal-matched typical talkers, as demonstrated by the comparison between the ASD group's MSEL age equivalence and chronological age of the TT and LT groups. Alternatively, the age differences also indicate that older children likely experienced additional exposure to verbs. Future work is needed to determine the exact factors that drove the

verb differences; the data analysed in the current study are insufficient to identify the exact mechanisms that might cause these differences. With regards to nouns, LTs showed lower proportions than TTs in the early vocabulary sizes (1-75 words). This is consistent with findings from MacRoy-Higgins, Shafer, Fahey, and Kaden (2016), who also found that late talkers had a lower percentage of nouns than age-matched and verbal-matched children. MacRoy-Higgins and colleagues suggested that the lower production of nouns in LTs might be an indication of a late-arriving vocabulary spurt. Interestingly, the proportion of nouns in children with ASD in our sample falls between the LTs and the TTs, something that cannot be explained by age since the ASD group has the largest language delay. This would mean in theory that the exact moment of the vocabulary spurt in children with ASD (relative to their vocabulary size) should be somewhere in between that of LTs and TTs (typically, a spurt has been observed once children acquire between 50 and 100 words; Bates et al., 1994). Nevertheless, the effect sizes for the noun differences are generally smaller than those for the verb differences.

Differences were found in four CDI semantic categories: Small Household Items, Animals, Toys and Vehicles. With respect to small household item words, the acquisition of the word "bottle" by TTs could have had a sufficient impact on the detected proportion differences due to the generally low production of small household items (0 to 2 words) at small vocabularies. The fact that the two oldest groups, LTs and children with ASD, differed from the youngest group, TTs, supports the argument that the differences identified are likely to be related to developmental discrepancies between the groups (i.e., age).

Curiously, we found differences in three CDI categories that are related to play: Animals (Real or Toy), Vehicles (Real or Toy), and Toys. Children typically produce a wider range of words that belong to these categories compared to Small Household Items (up to 5 words in the vocabulary sizes where differences were found). A noticeable difference between the type of toys that older children acquire compared to vocabulary-size-matched younger children is that they generally have features that can be manipulated where advanced fine motor skills are needed (e.g., vehicles, playdough). Parents of young toddlers might avoid giving their children toys with these characteristics,

which might represent a choking hazard, which potentially influences the type of words to be produced by the child. The differences in the proportion of animals produced in our sample follow the same pattern of differences in age, which support the age effect hypothesis. However, this is not the case for toy words and vehicle words, where the production of these words by children with ASD are somehow placed between TTs and LTs. These findings might suggest an association between word acquisition and play skills. The level of functional play in 3- to 5-years-old children with ASD has been found to be less elaborate and less diverse than vocabulary- and developmental-matched neurotypical children and children with Down syndrome (Williams, Reddy, & Costall, 2001). In a previous study, late talkers were also found to produced less sophisticated play, initiating fewer play scripts and producing more instances of non-functional play (Rescorla & Goossens, 1992). Our difference in toy-related words in our ASD and LT groups fits well with the prior literature documenting connections between play skills and language (e.g., Conner, Vance, Ryalls, & Friehe, 2014; Ingersoll & Schreibman, 2006). In any case, the effect sizes of these findings are small and suggest that any influence related to differences in development or to differences in play-skills are likely to have a weak impact on the semantic composition of the children's vocabularies.

Our study introduces some distinct methodological aspects in comparison to previous studies that might explain our findings. First, we differ from these studies in that we included late talkers as a third comparison group, allowing us to consider the potential effect of age or language delay. Second, we use a different vocabulary grouping system. Charman et al. (2003) examined vocabulary bins of many different sizes, starting with groups with small differences in the number of words produced (e.g., 1 to 5 words), up to groups with very large differences among children (e.g., +50 words). The aim of the Charman et al.'s (2003) arrangement was to facilitate the comparison of the CDI vocabularies of children with ASD with the normative sample collected by Bates et al. (1994), who also grouped children in this manner. Bates et al. (1994) claimed that early lexical development is characterised as being an 'unstable period'. In light of our results, it is perhaps the case that by grouping children in slightly larger vocabulary bins one might be able to

control for this predicted early variability (i.e., one bin of 1 to 25 words instead of three bins of 1 to 5, 6 to 10 and 11 to 20 words). Luyster et al. (2007) used no vocabulary grouping, and Rescorla and Safyer (2013) examined a group of children with very large vocabulary size differences (1 to 49 words, and 1 to 310 words). In addition to the methodological differences in vocabulary size bins, the two groups in the Ellis Weismer et al. (2011) sample (ASD vs LTs) differed in age. In comparison to our sample, the age gap between the ASD group and the LT group is much larger than that of Ellis Weismer et al.'s study (mean ASD: 30 months; mean LT: 25 months). Further, our ASD sample size nearly triples Ellis Weismer et al.'s ASD sample, and our late talker sample size is twelve times larger than their late talker sample.

2.4.3 Acquisition of High Social Verbs

Our analysis of social words found that children with ASD learned fewer highly social verbs than language-matched TTs and LTs with small verb vocabulary sizes. This finding may indicate that typically developing children may more reliably use social information to learn verbs. Previous research has suggested that verb learning is negatively influenced by a poor understanding of the speaker's social intentions in children with ASD (Parish-Morris, 2011). We contribute to this research by identifying social features not associated with the social interaction present at the moment of learning (i.e., adult speaking with the child), but in the words themselves, and that acquiring high-social verbs might be challenging to children who show difficulties in understanding social events. What seems to be contradictory is that, even though children with ASD showed a lower tendency to learn high social verbs, they managed to learn a higher proportion of verbs overall than their vocabulary-matched typically developing peers, suggesting that social features only have an influence on the type of verbs they acquire, not the quantity. This fact might be indicating an atypical use of verbs, which would be more directed to instrumental goals, rather than to social goals such as requesting for a joint attention activity or a coordinated and reciprocal play activity.

Since verbs associated with many event participants are harder to process by neurotypical children (Horvath, Rescorla, & Arunachalam, 2018), another possible explanation of why children with ASD produce less high

social verbs might be related to the difficulties they face with complex syntax, a characteristic typically observed in ASD (Tager-Flusberg & Joseph, 2003). Conversely, the number of social features in the first verbs learned by LTs resemble that of TTs, indicating that LTs are equally likely to attend to and learn verbs that are typically associated with interactions with other people. Late-talking toddlers have been previously found to socialise less than typically talking toddlers (Irwin, Carter, & Briggs-Gowan, 2002); however, our findings suggest that this may not negatively impact their learning of social verbs.

We also found that TTs with small verb vocabularies had more high-social verbs on average than TTs with larger vocabularies, indicating a preference for producing high-social verbs earlier. This is suggestive of a general social-word bias in early word acquisition, a word learning preference that, to our knowledge, no study has reported before. Although the visualization of our results suggests that LTs and children with ASD showed this social-word bias, our analysis determined that there were no significant differences between the two vocabulary sizes. The number of LTs and children with ASD in the large vocabulary groups were quite small, which provides us with low power to detect small effect sizes. Therefore, although children with ASD showed a reduced acquisition of high-social verbs at small vocabularies, we cannot discard the possibility that children with ASD have a weak social-word bias.

2.4.4 Limitations and Future Research

Although the current study has several strengths that allowed for novel insights into early lexical development in children with ASD, there are a number of limitations that must be kept in mind. The first limitation relates to the lack of information about the composition of the samples. Although we were able to gather information about ASD evaluation protocols used and nonverbal cognitive skills from the majority of the children with ASD, we were not able to gather this information for the children in the TT and LT groups. This means that there is a possibility of having cases of children in our TT and LT groups who could potentially be identified as having ASD at later times. This risk is especially relevant for the LT group. However, it should be noted that the incidence of late talkers is higher than that of ASD. The second

limitation is the criteria chosen to create semantic categories. We used the categories given by the CDI and treated the words within each category as semantically similar. Although this mirrored the method used in the prior literature, some special cases like sound effects could be categorised differently. Third, and most notably, we do not have data that provides sufficient insight into *why* lexical differences exist between the groups. Future experimental studies are needed to provide the necessary mechanistic accounts that explain the areas of distinction in the lexical profiles demonstrated by late talkers and children with ASD.

Future research should further investigate children's learning of verbs with varying social features in controlled learning situations to confirm our suggested interpretation related to children with ASD learning fewer high social verbs. Specifically, it would be of interest to determine whether the potential lower acquisition of high social verbs is related to the number of participants involved in the action or whether it is related to the degree of the social interest in the action. We failed to prove whether LTs and children with ASD show a social-word bias due to the small size of one of the comparison groups. Future research could confirm whether these children have a social-word bias at all by repeating the analysis with a larger sample of children with large verb vocabularies. In addition, our finding of the reduced noun-verb gaps in children with ASD and LTs motivate future research to further investigate the relation between this gap and language delay. Finally, based on our semantic category findings, future work should examine the relationship between specific play skills and word acquisition.

2.5 Conclusion

Although the proportion of words in the vocabularies of children with ASD is similar to neurotypical vocabulary-matched children in most semantic categories (supporting the dimensional account), the current study is the first to directly compare these three groups and to identify differences in two syntactic classes and four semantic categories. Most of the differences were found at small vocabularies and with small or medium impact on the composition of the children's vocabularies. In addition, the pattern of the group differences suggests age as a factor that drives most of the differences.

In addition to identifying similarities in many semantic categories, we also documented that LTs, TTs, and children with ASD demonstrate a noun bias; however, the degree of the noun-verb gap differed between the groups. We found that LTs and children with ASD had a smaller noun-verb gap relative to TTs, suggesting a link between language delay and noun-verb acquisition. Further, our results suggest that verb acquisition in children with ASD is influenced by the social features embedded in verbs, with these children primarily acquiring less-social verbs. However, more evidence is needed to confirm whether there is an absence of social-word bias or a weakened social-word learning bias in children with ASD. In sum, the current study has contributed to the ASD and LT literature by providing further information that highlights areas of overlap and distinction in early lexical development.

Chapter 3:
Semantic Maturation During the Comprehension-Expression Gap in Late and Typical Talkers

We examine how semantic maturation influences early lexical development by investigating the impact of contextual diversity—known to influence semantic development—on the comprehension-expression gap—the difference between the words children understand and those they produce. Study 1 compares semantic networks of the vocabularies of 3,685 typical and late talkers aged 16–30 months-old and finds that late talkers, with a longer preverbal phase, show lower contextual diversity in the nouns they produce, but higher contextual diversity for verbs. Study 2 compares computational network growth models of semantic maturation using Approximate Bayesian Computation (ABC) and finds that verbs require more semantic maturation than nouns, and typical talkers have more semantic maturation in the words they produce than late talkers.

3.1 Introduction

Toddlers understand more words than they can say. This is the well-known *comprehension-expression gap* (Goldin-Meadow, Seligman, & Gelman, 1976). When a child spontaneously produces a word for the first time (like "dog"), she already knows something about the meaning of that word (it is animated, furry, has a tail, etc.). However, what researchers understand less well is how necessary a *semantic maturation* (pre-productive learning of a word's meaning) is for word production. Moreover, what paths does this semantic maturation take for different populations of young learners? In particular, the comprehension-expression gap is larger for late talking toddlers (Thal & Bates,1988). This means that the words late talkers produce reside in their receptive vocabularies (i.e., comprehension) for longer. Could this mean that late talkers have the opportunity to form richer semantic representations of words before they produce them? Or is an atypical learning strategy by late

talkers responsible for creating weaker semantic representations, leading to delayed production?

One way to examine this is to focus on a well-known influence on early word learning, *contextual diversity* (Hills, Maouene, Maouene, Sheya, & Smith, 2009a; see also work on *semantic diversity*, Hoffman et al., 2013; Jones, Johns, & Recchia, 2012; Hsiao & Nation, 2018). Words with high contextual diversity (appearing with many other word types) are easier to learn by children (e.g., Rosa, Tapia, & Perea, 2017), presumably because contextual diversity enriches a word's semantic representation (Vergara-Martínez, Comesaña & Perea, 2017). In our example above, the child would enrich the concept of "dog" by encountering the word "dog" close to words like "furry" and "tail" within the speech. Although evidence shows that toddlers learn words with high contextual diversity earlier (e.g., Hills, Maouene, Maouene, Riordan, & Smith, 2010), late talkers have been hypothesised to produce more low-contextually-diverse words, leading to productive lexicons with weaker semantic links between words (Beckage, Smith, & Hills, 2011). Might this be a symptom of differences in the role of semantic maturation between typical and late talkers?

In this paper, we shed light on these questions by, first, examining the impact of contextual diversity on the productive vocabularies of a large sample of typical and late talkers, and second, by examining the performance of three computational models to disentangle the causes underlying different developmental pathways of semantic maturation.

Studies from psycholinguistics and cognitive science have identified distinct differences between production and comprehension (for a review, see Pickering & Garrod, 2013). Sahni and Rogers (2008; see also Stokes et al., 2019), examining both production and comprehension, found that productive vocabularies showed more influence of phonological density, whereas receptive vocabularies showed more influence of semantic density (which they based on shared perceptual features). Though the operational definition of semantics varies across studies, these results nonetheless suggest that the promotion of words from receptive to productive vocabularies is not based entirely on seniority (i.e., the length of time the word has been in the child's

receptive vocabulary before production). Some words in the receptive vocabulary become part of the productive vocabulary before words that have been in the receptive vocabulary for longer. If the factors influencing word promotion into and out of the receptive vocabulary differ across late and typical talkers, this should be reflected in the relative contextual diversity of the words in their productive vocabularies.

Contextual diversity is a measure of a word's context. Here we follow the approach of Adelman et al. (2006), and Jones, Johns, and Recchia (2012) in considering the word's contextual diversity a general feature of the word, measuring its likelihood of use in various contexts. Hills et al. (2009a) observed that the growth of early vocabularies clearly reflects a sensitivity to the contextual diversity in the structure of the language learning environment, outperforming frequency, phonology, and other word properties when predicting order of acquisition (see also Hills et al., 2010; Stella et al., 2018; Fourtassi, Bian, & Frank, 2019).

Focusing on how contextual diversity influences semantic maturation processes during lexical development might be crucial to explain early language delays, as late talkers, who do not present any disability or developmental disorder, usually show normal levels of comprehension despite exhibiting a delay in production (Thal, Marchman, & Tomblin referred to these children as 'late producers', 2013). Beckage, Smith, and Hills (2011) reported that typical talkers (TTs) and late talkers (LTs) show well-connected small-world semantic networks in their productive vocabularies, but late-talking children showed this to a lesser degree. Based on this, Beckage et al. proposed that LTs might be using an 'oddball strategy', which consists of a bias towards producing words that are less likely to co-occur with words that are already known, leading them to produce words with lower contextual diversity. This lower production of high contextual diversity words in LTs might indicate an atypical use of contextual diversity during semantic maturation.

Critically, different word classes have different patterns of contextual diversity, which suggests further evidence for differences in words' semantic maturation. Work of Gentner (1982) found that verbs have many tokens but few types in child-directed speech. This is in comparison with nouns, which have relatively fewer tokens but more

types. In the context of word co-occurrences, this may mean that verbs have more chances to form links with more word types than nouns, enhancing the contextual diversity of verbs. There is also evidence to suggest that contextual diversity is better correlated with age of acquisition for verbs than for nouns (Hills, 2013). If sensitivity to contextual diversity were similar for nouns and verbs, then theoretically verbs should be acquired earlier for their high contextual diversity, yet they are generally produced and understood later (Bates et al., 1994; Bergelson & Swingley, 2012). Of course, words have other characteristics that influence their learning, such as abstractness, which is higher in verbs and might explain their delayed onset (Hirsh-Pasek & Golinkoff, 2006; Bergelson & Swingley, 2012). Since word class may have an impact on a word's contextual diversity, by using a larger data set the present work offers an opportunity to extend the result of Beckage et al. (2011) to separate the analysis by word class. Further, contextual diversity facilitates the discovery of the semantic features of words. However, words like function words do not appear to have this advantage (e.g., Hills et al., 2010; Hills, 2013). Thus we focus on nouns and verbs here, which show the highest impact of contextual diversity. In addition, early lexicons are mostly composed of nouns and verbs (Bates et al., 1994).

3.2 The Present Studies

Here, we sought to answer the following research questions.

Q.1: *Does the productive vocabulary of LTs show a lower degree of contextual diversity than vocabulary-matched TTs?* In study 1, we answer this question by examining the contextual diversity of the nouns and verbs produced by a large sample of late and typically talking toddlers. In addition, we examine the network properties of the two talker groups. Based on Beckage et al.'s findings, we expect to find LTs' networks to show a lower degree of connectedness for both word types.

Q.2: *Do different word types go through different periods of semantic maturation before being produced?* We explore this question in Study 2 by comparing three computational models that tease apart the potential impact of a developing comprehension phase on posterior productive lexicons. Specifically, we are interested in determining whether verbs follow a different path of semantic maturation since verbs are presumed to have higher contextual diversity than nouns (a difference that we also test in our analysis).

Q.3: *Is contextual diversity a source of word information that mainly predominates in the preverbal phase (before the child produces her first word)?* Stella et al. (2018) suggested the existence of developmental changes in the sensitivity to contextual diversity during early stages of word learning. We address this question by letting the model change its sensitivity to contextual diversity during the preverbal and verbal phases of development.

Q.4: *Do LTs and TTs differ in the amount of semantic maturation necessary for words to move from comprehension to production?* We answer this question by examining the best free parameter values of our best computational model.

3.3 Study 1: Typical Talkers versus Late Talkers

3.3.1 Method

3.3.1.1 Sample, Identification of Late Talkers and Words for Analysis

A total of 5,520 vocabularies of American children aged 16 to 30 months were downloaded during April 2018 from *Wordbank* (Frank et al., 2017). 'Full Child-by-Word data' was selected under data, 'Words & Sentences' under forms, and '(American) English' under language. Various researchers contributed to gathering this data collected using the parental checklist MacArthur-Bates Communicative Development Inventory (CDI, Fenson et al., 1993). For the identification of LTs, we used the vocabulary norms provided by the CDI for children aged 16–30 months and classified children at or below the 20th percentile of vocabulary size for their age as LTs. This follows previous work (Beckage et al., 2011). For nouns, we included the following CDI categories: "Animals", "Vehicles", "Toys", "Food & drinks", "Clothing", "Body parts", "Household", "Furniture & rooms", "Outside" and "People". For verbs, we included all words categorised as "Action words". In order to measure contextual diversity and for clarity of interpretation, it was necessary to exclude homographs, e.g. "swing" (noun) and "swing" (verb); concepts composed of two words, e.g., "rocking chair"; and words that do not occur in the child-directed speech extracted from CHILDES (MacWhinney, 2000), which we describe in the next subsection. Our final word selection was composed of 286 nouns and 96 verbs. For a given word class (nouns or verbs), TTs were vocabulary-size matched to LT children. This necessarily excluded TTs with larger vocabulary sizes than our LT sample. We set a minimum of 10 words to avoid the high

variation produced by averaging small vocabularies. Productive noun vocabulary sizes ranged between 10 and 190 words; productive verb vocabulary sizes ranged between 10 and 70 words.

The final subset of children comprises 3,685 unique children, from which 3,211 were included in the noun analysis, and 1,949 children were included in the verb analysis. In the group of children for the noun analysis, 626 were identified as LTs; in the group of children for the verb analysis, 183 were identified as LTs. All children were aged 16 to 30 months (see Table 3.1). Some children with small vocabularies only produced nouns, and therefore they were included in the noun analysis only; similarly, children with large vocabularies and more than 190 nouns produced were excluded from the noun analysis, although they were included in the verb analysis since their verb vocabulary size was within the verb range analysed. A total of 1,475 children had both their verb and noun vocabularies analysed. In our noun analysis, within the LT group, 25.9% of children were female, 43.8% were male, and 30.4% had an unknown gender; within the TT group, 34.7% were female, 36.9% were male, and 28.4% had an unknown gender. In our verb analysis, within the LT group, 30.1% of children were female, 50.8% were male, and 19.1% had an unknown gender; within the TT group, 35.0% were female, 28.6% were male, and 36.4% had an unknown gender.

3.3.1.2 Contextual Diversity of Words and Network Analysis

To compute contextual diversity, we analysed American child-directed speech, taken from the CHILDES corpus (MacWhinney, 2000). We included all adults' speech directed to children up to 5 years old and computed the contextual diversity of each word. All children's speech was removed, and no free spaces were left between the adults' utterances. Following a surface proximity approach (see Evert, 2008), we determined the frequency in which each distinct word (node) in the corpus co-occurred with other words (collates). To do this, a matrix was populated by moving a window of size 10 word-by-word through the corpus. A window size of 10 was selected since it best predicted age of acquisition (see model details in Appendix C). The word at the start of the window was used to index the row [i,], and any word encountered downstream and within the window of the starting word was used to index the column [, j]. When two words co-occurred, a value of one was added to position [i, j] in the matrix. The resulting weighted matrix was transformed into a binary matrix measuring contextual diversity with word-types by setting [i, j] =1 for all [i, j] > 0. Finally, we

extracted from this matrix a smaller matrix with nouns and verbs only. We then calculated the contextual diversity value for each word by adding the sum of the row and the sum of the column of this submatrix. Therefore, the contextual diversity value of a word reflects the number of semantic links that the word in question has with other verbs and nouns in our sample.

For the network analysis, we used two submatrices—one each for nouns and verbs—that we extracted from the matrix of co-occurrences described above. Words in the child's lexicon are represented as nodes, and the edges between nodes indicate semantic relatedness, inferred from the co-occurrence of each word with all other words within the speech stream. Network analysis in cognitive psychology (see Vitevitch, 2019) has been extremely successful in detecting structural differences in language acquisition (Bilson, Yoshida, Tran, Woods, & Hills, 2015; Hills, Maouene, Maouene, Sheya, & Smith, 2009b) and lexical processing (Vitevitch, Chan, & Goldstein, 2014; Vitevitch, Ercal, & Adagarla, 2011). Network statistics were computed using R and the *igraph* package, version 1.0.1 (Csárdi & Nepusz, 2006). Once all the words were connected in each vocabulary using the binary matrix described above, the clustering coefficient and average path length were calculated for each child's undirected network. Clustering coefficient measures the degree to which nodes in a network tend to cluster together. Specifically, we calculated the local clustering coefficient, which measures the number of links that the neighbours of a node have among themselves. Average path length is the mean shortest path between all pairs of words in a network, describing the level of global access.

We conducted correlational analysis of the following word features: contextual diversity, frequency, concreteness, and word length. For this, we computed word frequency from our CHILDES sample, calculated word length (phonemes), and assigned concreteness values to each word in the analysis, which were taken from Brysbaert, Warriner and Kuperman (2014).

Table 3.1 Distribution of Typical Talkers (TT) and Late Talkers (LT) across Productive Vocabulary Sizes

Vocabulary size	Group				Total	
	Typical talkers		Late talkers			
Nouns	n	age *M* [range]	n	age *M* [range]	n	age *M* [range]
(9,29]	519	16.6 [16, 20]	272	20.9 [16, 30]	791	18.1 [16, 30]
(29,49]	383	17.7 [16, 23]	89	24.2 [20, 30]	472	18.9 [16, 30]
(49,69]	274	18.8 [16, 26]	68	25.5 [23, 30]	342	20.1 [16, 30]
(69,89]	227	19.8 [16, 27]	68	26.5 [24, 30]	295	21.4 [16, 30]
(89,109]	213	20.9 [16, 27]	40	27.3 [25, 30]	253	21.9 [16, 30]
(109,129]	219	21.4 [16, 28]	36	28.4 [27, 30]	255	22.4 [16, 30]
(129,149]	216	22.1 [16, 28]	25	28.6 [27, 30]	241	22.8 [16, 30]
(149,169]	244	23.8 [16, 30]	23	29.1 [27, 30]	267	24.2 [16, 30]
(169,190]	290	24.5 [16, 30]	5	29.4 [29, 30]	295	24.6 [16, 30]
Total	2585	20.1 [16, 30]	626	24.0 [16, 30]	3211	20.8 [16, 30]
Verbs						
(9,29]	698	20.9 [16, 29]	135	27.1 [17, 30]	833	20.9 [16, 30]
(29,49]	509	23.8 [16, 30]	41	28.4 [20, 30]	550	23.8 [16, 30]
(49,70]	559	25.2 [16, 30]	7	29.0 [28, 30]	566	25.2 [16, 30]
Total	1766	23.1 [16, 30]	183	27.5 [16, 30]	1949	23.5 [16, 30]

Statistical analyses used generalised additive models (GAM) using the *mgcv* package and the *bam* functions in R (see Wood, 2011). GAM was chosen over simpler models because, first, the dependent variables of interest (average contextual diversity, clustering coefficient, and average path length) were highly correlated with vocabulary size and the talker groups differed in the number of words produced; second, homogeneity of variance was violated, resulting in the need to randomize vocabulary size. In our GAMs, the independent variable talker type was set as a fixed term, and vocabulary size was set as a smooth term. Statistical assumptions were verified using the *gam.check* function. The Akaike information criterion (AIC) of the models presented here were significantly higher than the AICs of the same models without talker type as a fixed term. An interaction term between vocabulary size and type of talker did not significantly reduce the models' AIC.

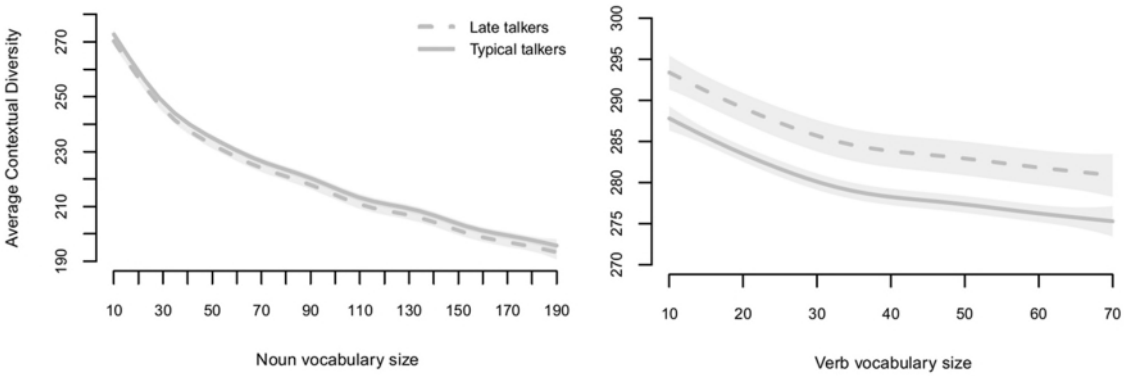
3.3.2 Results

3.3.2.1 Word Characteristics

Following Gentner (1982) and justifying the distinction between verbs and nouns we make in our analysis below, verbs present higher contextual diversity than nouns (verbs: *Mdn* =268, nouns: *Mdn* =170.50, $W = 5854.50$, $p < .001$, $d = .95$, 95% CI [−111.00, −72.00]). Further, in our best multilevel binomial regression model (AICs compared), contextual diversity significantly predicted word class ($b = 1.35$, $z = 5.58$, $p < .001$) in a model where frequency was randomised. We also computed the contextual diversity of every unique word in the CHILDES corpus, and our correlational analysis showed that, as Hills et al. reported in their work in 2010, high contextual diversity words are produced earlier ($r = 0.46$; See Appendix D for details).

Contextual diversity is positively correlated with word frequency ($r(369) = .47$, $p < .001$), and negatively correlated with concreteness ($r(369) = -.43$, $p < .001$). Concreteness and frequency are negatively correlated ($r(369) = .45$, $p < .001$). Word length is not correlated with either frequency or contextual diversity (concreteness: $r(369) = -.0008$, $p > .05$; contextual diversity: $r(369) = -.005$, $p > .05$), but it is weakly correlated with concreteness ($r(369) = .10$, $p < .05$))

Figure 3.1 Average Contextual Diversity in the Noun (Left Panel) and Verb (Right Panel) Vocabularies of Late Talkers and Typical Talkers across Vocabulary Sizes.



Note. Shadows around the curves represent 95% confidence intervals.

3.3.2.2 Contextual Diversity

As shown in Figure 3.1, LT and TT children differ in the contextual diversity of the nouns and verbs they produce (fitted values for nouns and verbs comes from two separate models). Regarding nouns, LTs produced lower average contextual diversity than typical talkers, adjusted $R^2 = .80$, $F(1, 3201.23) = 20.54$, $p < .001$, 95% CI [1.35, 3.49]. In the case of verbs, LTs produced higher average contextual diversity in their verb vocabularies than TTs, adjusted $R^2 = .13$, $F(1, 1943.91) = 35.08$, $p < .001$, 95% CI [−7.48, −3.70]. In addition, both types of talkers show a decrease in contextual diversity as they produce more words. This is true for nouns, $F(7.77, 8.60) = 1385$, $p < .001$, as well as for verbs, $F(3.09, 3.83) = 53.57$, $p < .001$.

Since there is a gender imbalance between our talker groups, there is a chance that the differences identified above are related to the higher proportion of male participants in the LT group. This is because female children are known to have more advanced early language skills than male children (e.g., Eriksson et al., 2012). We re-ran our analysis with males only, and also excluded children whose mother's education was below college level since maternal education was found to be a reliable predictor of children's language (e.g., Reilly et al., 2007; Reilly et al., 2010) (Nouns TT: $n = 541$, age in months $M = 21.0$, age $SD = 3.4$, age range = 16-30; Nouns LT: $n = 165$, age $M = 23.6$, age $SD=3.9$, age range =16-30; Verbs TT: $n = 429$, age $M = 23.1$, age $SD = 3.6$, age range = 16-30; Verbs LT: $n = 41$, age $M = 27.5$, age $SD=2.4$, age range = 20-30).

The same differences emerged between the groups in the noun analysis, with LTs showing lower contextual diversity averages in their vocabularies than TTs, adjusted R^2 = .79, $F(1, 697.9544) = 19.75$, $p < .001$, 95% CI [2.67, 7.12]. Similarly, LTs showed higher contextual diversity in their verb vocabularies than TTs, adjusted $R^2 = .10$, $F(1, 466.9994)$ 15.22, $p < .001$, 95% CI [-12.00, -3.87].

3.3.2.3 Network Properties

The results of the network analyses also indicate differences between LTs and TTs. With regards to nouns, LTs exhibited lower clustering coefficient, adjusted $R^2 = .79$, $F(1, 3202.10) = 18.73$, $p < .001$, 95% CI [0.0029, 0.0078], and higher average path length compared to TTs, adjusted $R^2 = .82$, $F(1, 3201.93) = 32.05$, $p < .001$, 95% CI [−0.020, −0.097]. With regards to verbs, LTs presented higher clustering coefficient, adjusted $R^2 = .22$, $F(1, 1945.19) = 24.25$, $p < .001$, 95% CI [−0.012, −0.005]) and lower average path length than TTs, $F(1, 1946.00) = 22.72$, $p < .001$, adjusted $R^2 = .21$, 95% CI [−0.007, −0.017]. In sum, both contextual diversity and network analysis of LT and TT indicate differences in their underlying semantic maturation driving noun and verb production. In Study 2, we use computational models to better understand the plausible role of semantic maturation in producing these results.

3.4 Study 2: Comprehension-Expression Models

The differences between nouns and verbs and typical and late talkers offer a unique opportunity to investigate a potential role for semantic maturation in word development. To do this, we developed a series of models that each test specific hypotheses about word development. Currently, the best existing semantic network model is preferential acquisition, introduced by Hills et al. (2009a, 2010), which performs best when its learning rule is based on contextual diversity, as opposed to, for example, frequency, phonology, or perceptual features (e.g., see Sailor, 2013; Bilson et al., 2015; Shrestha et al., 2015; Stella, Beckage, & Brede, 2018). In what follows, we first describe the preferential acquisition model and then introduce a series of extensions to evaluate the following: a) semantic maturation in the pre-verbal phase (Progressive Preferential Acquisition), and b) feedback from productive vocabulary to comprehension (Progressive Lure of the Associates). These two models further allow us to evaluate c) differences in semantic maturation between nouns and verbs, and d) differences in growth rules between preverbal and verbal development. For

completeness, we also verified that models based on frequency were outperformed by contextual diversity, which, following prior work (e.g., Hills et al., 2009a, 2010), is the learning cue we use here.

3.4.1 Method

3.4.1.1 Preferential Acquisition Model

Preferential acquisition (PA) proposes that children leverage contextual diversity in the language learning environment to learn new words. A word *i* is selected to be learned out of a pool of words to be learned *W* (nouns and verbs) with a probability proportional to its contextual diversity in the learning environment:

$$P(i) = \frac{(K_i+1)^{\beta}}{\sum_{i \in W}(K_i+1)^{\beta}} \quad (1)$$

The probability of word *i* being selected depends on its contextual diversity value K_i, computed as described in Study 1. The denominator sums these values over all possible words, so the total probability sums to 1. Sensitivity to contextual diversity, represented by the β in Figure 3.2, is fit to the data and is the primary factor influencing how words move from the learning environment into the productive vocabulary. When β is greater than 0 there is a preference to add more contextually-diverse words to the network. When β is 0, all words are treated equally with respect to contextual diversity. Words are chosen based on equation (1), one word at a time until no unknown words remain to be learned.

3.4.1.2 Progressive Preferential Acquisition Model

In the Progressive Preferential Acquisition (PPA) model, word choice follows the PA model, but words move first into a comprehension layer. Each time a word is chosen using equation (1), its value in the comprehension layer receives a maturational *boost,* increasing the strength of the words' semantic representation in the comprehension network. We used a value of 0.5 for *boost*, but a value of 1 produced similar results. When a word matures above a threshold in the comprehension network, represented by the τ in Figure 3.2, it then moves into the productive vocabulary. Words are sampled using equation (1) until no unknown words remain to be learned. PPA

adapts to the PA framework the maturational proposal in the prior simulation work of McMurray (2007) and Nematzadeh, Fazly and Stevenson (2014).

3.4.1.3 The Progressive Lure of Associates Model

The progressive lure of the associates (PLA) extends the PPA model one step further by introducing a mechanism for known words to enhance comprehension of related words. This follows the logic of the lure of the associates model (Hills et al., 2009a), which further leverages the principle of mutual exclusivity underlying preferential acquisition by allowing known words that appear together with unknown words to further facilitate their acquisition. Hills et al. (2009a, 2010) found roughly equivocal performance between preferential acquisition and lure of associates, both outperforming a variety of other models. The additional data and differences between typical and late talkers and nouns and verbs in our present analysis allow us to further tease apart these model differences.

The progressive lure of the associates model (PLA) assumes that words move into comprehension and production exactly as defined for the PPA model. However, PLA adds an additional mechanism for boosting words, such that words associated or linked to a newly produced word gain an additional *boost* to their comprehension. This is represented for the PLA model in Figure 3.2 by the arrow from production to comprehension and by the dotted arrows in the comprehension network radiating out from produced words to near associates.

In the base versions of PPA and PLA, we allow β to change dynamically between the preverbal, β_p, and verbal, β_v, phase, such that $\beta_p \neq \beta_v$. In addition, the threshold for nouns, τ_n, and verbs, τ_v, is allowed to vary independently as well, such that $\tau_n \neq \tau_v$.

3.4.1.4 Extending the Models

The base PA, PPA, and PLA models offer a framework for further examining additional hypotheses about developmental changes in semantic maturation. Furthermore, they also allow to test some of their underlying assumptions. In particular, we can test the validity of the dynamic change in sensitivity to contextual diversity underlying semantic maturation between preverbal and verbal development, as well as differences in semantic maturation between nouns and verbs. We do this as follows.

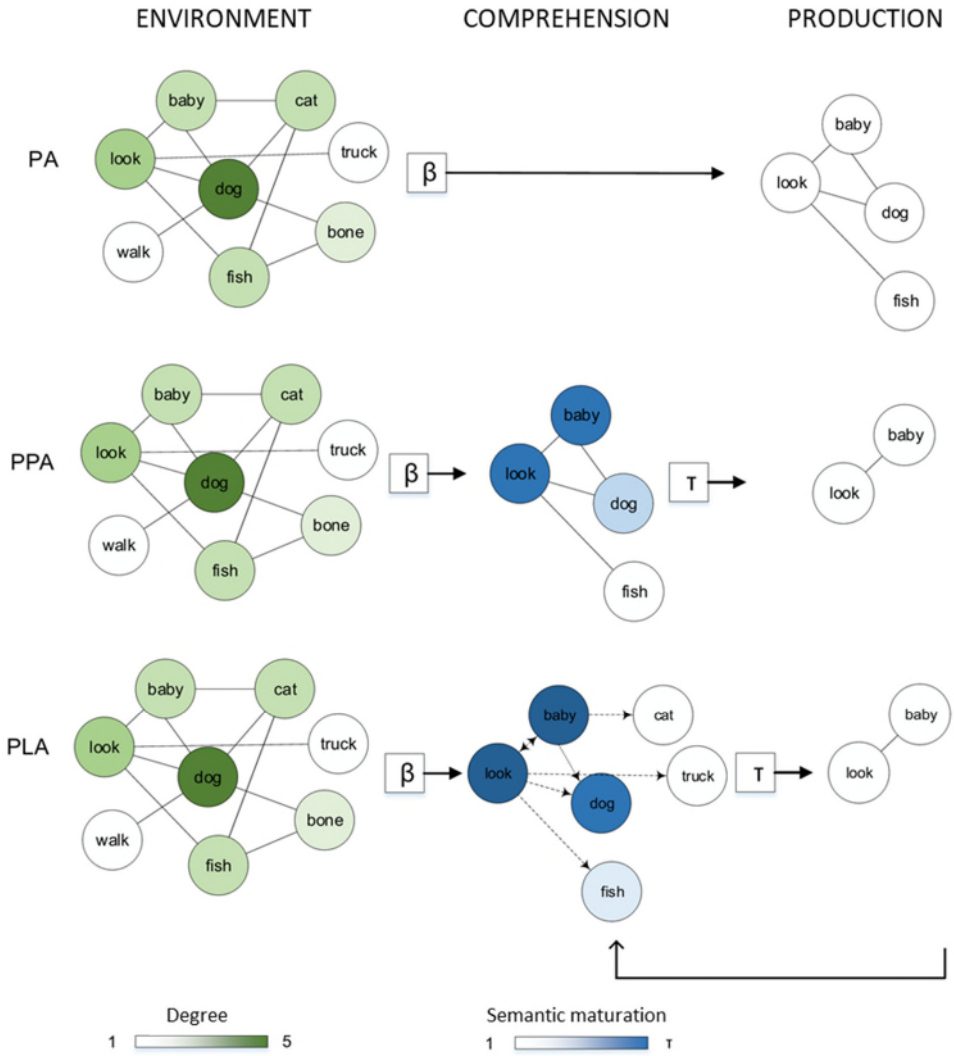
Changes between preverbal and verbal development. Recent work suggests there may be developmental changes in sensitivity to contextual diversity during early word learning (Stella et al., 2018). To test this, we compare the dynamic $\beta_p \neq \beta_v$ models described above (PPA^0 and PLA^0) with models that use a single, β, such that $\beta_p = \beta_v$ (PPA^1 and PLA^1), which assume that the sensitivity to contextual diversity is the same during preverbal and verbal development.

Differences in semantic maturation between nouns and verbs. To evaluate the potential for differences in semantic maturation between nouns and verbs, we extend the base models described above, for which $\tau_n \neq \tau_v$, by comparing it with a single τ model (PPA^2 and PLA^2), which assumes that nouns and verbs mature at a similar rate.

3.4.1.5 Parameter Estimation and Model Comparison

Though maximum likelihood models were initially described for parameter estimation and model comparison of the generative network growth model underlying PA (see Hills et al., 2010), similar analytical solutions for PPA and PLA are computationally intractable due to their complicated dependency structures (see Hills et al., 2010). We solve this problem using two approaches that trade-off search efficiency (grid search) versus model complexity penalization (Approximate Bayesian Computation). Both of these methods confirm the findings of the other regarding model comparisons. Both models are also approximations with respect to parameter estimation—a common challenge for rugged high-dimensional landscapes. Our aim here is to capture the qualitative switch between nouns and verbs in LT and TT populations, while also providing the best quantitative fit to the data.

Figure 3.2 The Three Primary Models Introducing Semantic Maturation in a Comprehension Layer and Feedback from Production to Comprehension.



Note. The three primary models introducing semantic maturation in a comprehension layer and feedback from production to comprehension. From top to bottom: Preferential Acquisition (PA), Progressive Preferential Acquisition (PPA), and Progressive Lure of Associates (PLA). For each model, words are sampled from the learning environment according to equation (1) with sensitivity parameter (β). After each sample, the word is either added to production (PA) or its semantic maturation is boosted in the comprehension network (PPA and PLA). In PPA and PLA, when words exceed a maturation threshold (τ) in the comprehension network, they move into the productive vocabulary. In the PLA model, movement into the productive vocabulary leads to an additional boost for associates in the comprehension network. Contextual diversity is synonymous here with the number of links each word has with other words in the learning environment.

To identify the best fitting parameters, we used a grid search, which provided uniform coverage of the search space. We used 500 vocabulary growth simulations for each set of parameters and for each model, with the optimal parameters identified as those that minimised the mean squared error (MSE) between the observed and simulated data. To compute the MSE, we split the simulated ordered vocabulary into a noun ordering and a verb ordering (following Figure 3.1). The average contextual diversity was computed for the words in each vocabulary size network, creating a vector, V_i, of contextual diversity scores corresponding to a single developmental trajectory. The mean of these trajectories, $\bar{V}$, was computed for the 500 simulations and compared with a similar vector for the observed data, V_o, which was computed using the data presented in Study 1. The optimal parameters that minimised the MSE between $\bar{V}$ and V_o are presented in Table 3.3.

Though the comparison of the MSE between models is consistent with the model comparison we describe below, MSE does not take into account model complexity. Thus we turn to Approximate Bayesian Computation (ABC), which is well-adapted for complex model comparisons where maximum likelihoods are unavailable, and which has growing popularity in fields such as evolutionary biology and genetics (e.g., Fagundes et al., 2007; Fraïsse al, 2018; Roux et al., 2014). ABC randomly samples model parameters from prior distributions and computes posterior likelihoods via model simulation by comparing simulated and observed data using summary statistics (Hartig et al., 2011; see also the *abc* R package, Csillery, Francois, & Blum, 2012). For each model's free parameters, we used uniform priors set to bound the optimal parameter values identified using the grid search described above. Priors distributions were identical across models: β_p = [0,2]; β_v = [0,1]; τ_n and τ_v[2,14]. We then randomly sampled a set of parameters from the prior distributions and used them to simulate vocabulary growth trajectories, iteratively sampling from the pool of available words (*n*=382). Following the suggestion of Beaumont (2004) regarding reducing summary statistic complexity, we averaged across progressive size networks for each simulated growth trajectory in bins of size 20 to produce a simulated growth vector for that parameter set, $\bar{V_s}$, and a corresponding vector for the observed data, V_o. In total, we simulated a total of 100,000 simulations for each model.

The posterior distribution is composed of those samples for which the Euclidian distance between the simulated and observed summary statistics, $\rho(V_s, V_o) = |V_s - V_o|$, is below a threshold tolerance, formally: $\rho(V_s, V_o) < \epsilon$. Accepted parameters approximate the posterior distribution (Hartig et al., 2011). The model comparison and Bayes factors were calculated based on the posterior distributions using the 'abc' R package using the local regression method, with similar results for the best model using the neural network method (see Csillery et al., 2012 for details).

Table 3.2 Posterior Probabilities of The Base Models and their Extensions.

	Model						
Talker Type	PA	PPA[0]	PPA[1]	PPA[2]	PLA[0]	PLA[1]	PLA[2]
TT	0.0000	0.4988	0.0324	0.3772	0.0438	0.0091	0.0387
LT	0.0000	0.6601	0.0017	0.3257	0.0061	0.0001	0.0063

Note. LT= late talkers; TT= typical talkers (TT). PA= preferential acquisition model; PPA= progressive preferential acquisition model; PLA= progressive lure of associates model. ϵ is 0.10; similar results are found with ϵ = 0.20 and ϵ = 0.30.

[0]: $\beta_p \neq \beta_v \; \tau_n \neq \tau_v$;

[1] : $\beta_p = \beta_v \; \tau_n \neq \tau_v$;

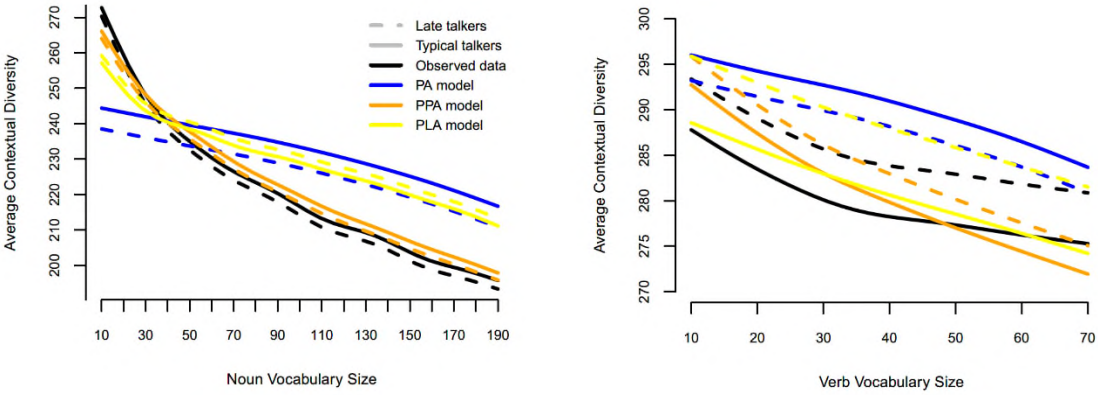
[2] : $\beta_p \neq \beta_v \; \tau_n = \tau_v$;

Table 3.3 Best Parameter Values for Each of the Best Fitting Base Models.

Parameter	Models					
	PA		PPA		PLA	
	LT	TT	LT	TT	LT	TT
β_p			0.75	0.75	1.3	1.3
β_v	2.02	1.69	0	0	0.1	0.2
τ_n			6.5	7	2.5	1.5
τ_v			10	18	3	8.5
MSE	492.1	355.5	364.9	229.5	447.6	329.02
Model Probability	0.0000	0.0336	0.9902	0.8860	0.0098	0.0804

Note. LT= late talkers; TT= typical talkers (TT). The best parameter values for each base model was calculated after minimizing MSE. Approximate Bayesian Computation was used to compute models' posterior probabilities. ϵ is 0.10; similar results are found with ϵ = 0.05, ϵ = 0.20 and ϵ = 0.30.

Figure 3.3 The Predicted Average Contextual Diversity of the Noun and Verb Vocabularies of TT and LT Populations as Modelled by the Three Base Computational Models.



Note. PA: Preferential Acquisition; PPA: Progressive Preferential Acquisition; PLA: Preferential Lure of Associates. Best parameter values calculated for each base model were used for each talker type. All models are compared against the observed data, which is in black. Each model is represented by a colour and each type of talker is represented by a type of line.

3.4.2 Results

Table 3.2 displays the posterior probabilities of all the models in comparison with one another split by talker type. The posterior probabilities indicate that the base PPA has the highest likelihood to have generated both the LT and TT target observed data. For completeness, the best parameter values for each of the three base models are shown in Table 3.3, including the most probable PPA. Critically, for the PPA model, the attention parameters are not different between the LT and TT populations. However, semantic maturation thresholds for nouns and verbs are different from one another, and also different between talker type. LTs have a smaller gap between verbs and nouns than TTs, which leads LTs to start adding verbs to production soon after they start producing nouns.

We simulated the vocabulary growth of the three base models and visualised the results in Figure 3.3. We produced 500 networks for each model using the best parameters in Table 3.3 for each talker type. Then, we averaged the contextual diversity of each network at each vocabulary size, separately for nouns and verbs. As can be seen in Figure 3.3, the PPA model (orange lines) is the only one that exhibits a similar growth curve to what was observed in children's vocabularies in both nouns and verbs (black

lines), and it also captures the qualitative shift in contextual diversity between nouns and verbs found in Study 1.

3.5 Discussion

The present results demonstrate that comparison of the comprehension-expression gap between LT and TT populations provides a useful entry point for examining the potential role of semantic maturation in word development. Our best performing model, the progressive preferential acquisition model, suggests that words that enter the receptive vocabulary experience a period of semantic maturation before moving into production. Nouns and verbs mature at different rates. In addition, contextual diversity appears to be most important during the preverbal phase, with the sensitivity to contextual diversity falling to zero during the verbal phase. This suggests that preferential acquisition is primarily a strategy of early comprehension.

Although our observed data indicate differences in the role of contextual diversity between LTs and TTs depending on the type of word, the PPA model allowed us to tease apart this relationship. The results suggest that it is not that LTs pay more attention to contextual diversity for verbs, but rather that verbs require less semantic maturation for LTs than for TTs. The similar attention parameters between LTs and TTs corroborates Vuksanovic and Bjekic's findings (2013), which found that LTs have a comparable frequency of joint attention as TTs in word learning situations. Further, our findings are consistent with those of McGregor, Newman and Reilly (2002) who found that children with Developmental Language Disorder (DLD) showed poorer semantic knowledge about words already produced. Thus, LTs are delayed off the starting block, but once they start using contextual diversity to learn words, they require less information before moving words into production, giving rise to the signature pattern of noun and verb development we see here.

What factors influence these differences in semantic maturation is still an open question. These may be related to children's abilities or the learning environment, or a combination of both. The variation among families in the quantity and quality of child-directed speech is known to predict children's vocabulary growth (Hart & Risley, 1995; Hoff & Naigles, 2002). This is an important issue, as the differences between LTs and TTs might be related to differences in the degree of contextual diversity in parental speech, and not so much on children's abilities to exploit contextual diversity as a word-

learning strategy. Late talkers can use cross-situational evidence to learn words in a word learning intervention (Alt et al., 2014) and thus future work needs to look more closely at how the learning environment for individual children predicts which words they learn. This is the focus of ongoing work in numerous laboratories, including our own. The insight into semantic maturation from the PPA model provides an advance in computational approaches towards achieving that.

Though contextual diversity is one of the stronger predictors of early word learning, numerous other word features have been shown to be at play, such as phonological cues, perceptual features of the objects words refer to, concreteness, and other semantic features (e.g., Engelthaler & Hills, 2017; Storkel & Lee, 2011; Stokes et al., 2019). The results of the present work suggest these factors may be especially important after the preverbal phase, where contextual diversity appears to play less of a role. In this sense, contextual diversity may set the stage for future development, but may not remain an active player in word learning once children become increasingly active in their language learning. Future modelling work on word maturation should aim to incorporate the potentially dynamic roles of attention to various cues in the learning environment, especially phonological influences on word promotion from comprehension to production (Sahni & Rogers, 2008; Stokes et al., 2019; and see Stella et al., 2018). In sum, our findings suggest that semantic maturation during the comprehension-expression gap is driven by contextual diversity, and differences in the threshold for semantic maturation help to explain the differences between typical and late talkers as well as the differences between nouns and verbs. Given the potential influence of contextual diversity in promoting words from the receptive to productive vocabulary, practitioners could target words with high contextual diversity for vocabulary interventions and guide parents in how to introduce these words in different natural situations/contexts.

Chapter 4: Individual Differences in the Semantic Structure of Child-Directed Speech

The present study investigates the relation between language environment and language delay in 63 British-English speaking children (19 typical talkers (TT), 22 late talkers (LT), and 22 late bloomers(LB)) aged 13 to 18 months. Families recorded daily routines and marked the new words their child produced over a period of six months. The language environment experienced by the three talker groups differed in their structural-network properties, with LT environments being less semantically rich and less well-connected. The semantic associations between words learned in LT environments led LTs' expressive vocabulary to have weaker network structures. Further, LTs understood a different set of words which reflected the characteristics of their environments. We discuss the potential consequences that differences in parental speech might have on language delay and language processing.

4.1 Introduction

The semantic structure of natural speech has been shown to facilitate language acquisition as well as to provide a way in which humans organise their lexicons for efficient language processing (e.g., Hills, Maouene, Maouene, Riordan & Smith, 2010; Steyvers & Tenenbaum, 2005; Steyvers, Shiffirin, & Nelson, 2005). Yet, it is not well understood how individual differences in the language environment might be associated with individual differences in language acquisition and lexical processing in early childhood. Regarding language acquisition, research in adults and children has suggested that the higher the number of linguistic contexts in which a word appears, the easier the word is to learn (e.g, Rosa, Tapia, & Perea, 2007; Pagán & Nation, 2019). This important distributional characteristic in language is known as *contextual diversity*, which has been suggested to support cross-situational learning (e.g., Redington, Chater & Finch, 1998; Mintz, Newport, & Bever, 2002), and to aid the enrichment of the words' semantic mental representation (Hills, 2013; Vergara-Martínez, & Perea, 2017). Despite the importance of contextual diversity in lexical learning, no research has evaluated the

impact that different intensities of this environmental feature have on early language development, particularly on whether semantically poor language environments might be associated with low rates of lexical development in early childhood (i.e., early language delay).

With regards to language processing, network properties of human lexicons have explained the behaviour observed in different lexical tasks such as word recognition (e.g., Chan and Vitevitch, 2009) and word retrieval (e.g., Goldstein and Vitevitch, 2014). In language acquisition studies, the way in which words co-occur with other words in child-directed speech has been used to derive semantic relatedness between words (Jones & Mewhort, 2007; Collins & Quillian, 1969). These semantic relations between words are used to estimate how words in the children's lexical networks are connected with each other. Importantly, this approach underscores the important role of the language environment in how children organise their lexicons, which theoretically influences the way in which they process language. Previous work has suggested weaker language processing in late-talking toddlers based on the type of words that these children learn, which lead to less well-connected vocabularies (Beckage, Smiths, & Hills, 2012). A major assumption in Beckage's et al.'s work is that late talkers are exposed to the same linguistic environment as typical talkers, and therefore, the same semantic relations between words are presumed to be learned by all children in the sample. However, differences in the language structure of individual families are naturally expected, and so any conclusion on children's language processing based on environmental measures would need to take into account the children's unique linguistic environments.

The present work analyses the network properties of the vocabularies of 63 children and that of the child-directed speech of their respective families. Their vocabularies and environments were remotely tracked for a period of 6 months, allowing us to collect a large amount of data about their language learning environments as well as to identify which children became late bloomers at the end of the study. This unique dataset allowed us to investigate the potential influence that individual differences in the semantic structure of the environment can have on early language acquisition. Before we describe our study in more detail, we first describe the strong evidence for contextual diversity, followed by the description of what is known about the late-talking population in relation to how they learn and process language and their learning context.

4.1.1 Language Acquisition and Contextual Diversity

In hearing a new word, children face the challenge of disambiguating its meaning from the many potential references present in the same scene. When the same word is heard in different scenes the word-referent remains constant across situations, aiding the infant to solve the problem of indeterminacy (Quine 2013; Hills, 2013). In addition to this cross-situational learning strategy, if the child knows the word for some referents in the scene, she could accelerate the word-reference mapping using the principle of mutual exclusivity (one word only refers to one referent). However, this is not the only advantage of encountering words in different contexts. Words that appear in multiple distinct situations also provide the advantage of interacting with other words and offer the opportunity to extract important semantical properties of the words. For instance, the word *duck* is likely to occur in bath time situations, which motivates the infant to make semantic connections of the word *duck* with other bath-related words like *bubble*, *towel* or *splash*; similarly, the word *duck* is likely to happen in playtime sessions, together with other animal words, like *horse* or *pig*. In these instances, the fact that the word *duck* appears across two very different situations encourages the child to place the word *duck* in two different mental semantic categories: animal words and bath-related words.

The contextual diversity of words has been used to successfully predict their age of acquisition, with high contextual diversity words having a high likelihood to appear in the early stages of vocabulary development (Hills, Maouene, Maouene, Riordan & Smith, 2010). Various studies have shown that words learned in many different contexts are easier to process and recognised faster (Goldinger & Azuma, 2004; McDonald & Shillcock, 2001; Nelson & Shiffrin, 2006; Pexman et al., 2008; Hsiao & Nation, 2018; Pagán, Bird, Hsiao & Nation 2020). Further, contextual diversity was found to be better than frequency in predicting reaction times in word naming and lexical decisions tasks (Adelman, Brown & Quesada, 2006).

4.1.2 Children with Early Language Delay and their Vocabulary Structure

The majority of children with an early language delay do not show any disability or developmental disorder that explains that delay. Most late talkers accelerate their word learning rate until they catch up with their same-age peers during their first years; however, many of these late bloomers experience future delays in specific language

abilities or language-related tasks, such as in understanding and producing complex sentences at age five (Rescorla & Turner, 2015) and in non-word repetition tasks at age 11 (Conti-Ramsden, Botting, & Faragher, 2001). Multiple studies have shown that late talking toddlers exhibit atypical learning mechanisms (for a review see Desmarais et al., 2008), which motivated several studies to examine the lexical profiles of late-talking toddlers in search of differences compared to typically-developing toddlers (Jiménez, Haebig, & Hills, 2020; Ellis Weismer 2017; Rescorla 2009). In a recent study by Jiménez and colleagues (2020) late talkers were found to produce higher proportions of verbs and lower proportions of nouns than vocabulary-matched typical talkers, which led late talkers to present a weaker noun bias. Other studies found that late talkers produced fewer intransitive and ditransitive verbs (Olswang et al. 1997) and fewer manner verbs (Horvath et al. 2019) compared to same-age children.

The network structure of the vocabularies of late talkers seems to be influenced by the type of words these children produce. Beckage and colleagues (2011) showed that the semantic structure of late-talking toddlers' vocabularies was found to be less clustered and generally less well-connected compared to vocabulary-matched typical talkers, suggesting an impoverished use of a contextual diversity strategy. In cognitive network science, words with low phonological clustering coefficient are recognised faster than words with high clustering coefficient (Chan & Vitevitch, 2009); in contrast, words with high clustering coefficient have the advantage to be easier to learn (Goldstein and Vitevitch, 2014). Further, words that bridge cluster of words in the human lexicon are suggested to play an important role in processing language (Ferrer i Cancho & Solé, 2001). Words that are highly-connected in language are also likely to be produced as a target in free-association tasks (Hills et al., 2010). Toddlers are better at retrieving newly learned words that were paired with their semantic characteristics (shape or function), meaning that words with many semantic links (high degree) facilitate word learning (Capone & McGreggor, 2005). Given the relation between the connectivity of lexical network's and lexical processing, the examination of the network properties of children's vocabularies should be informative of internal linguistic processes.

Lexical processing in young children has been typically operationalized by the eye-tracking patterns that children produce during word recognition tasks. In a longitudinal study by Fernald and Marchman (2012), late talkers that remained delayed showed a lower proportion of time fixating on the target object in a word recognition task than

those late talkers that later on bloomed. In another study by Ellis and colleagues (2015) the authors utilised more fine-grained measures and found that late talkers' patterns of looking to be qualitatively different from that of same-age typical talkers. One important question arising at this point would be to what extent the ability to process language in young children is closely related to the semantic connectivity in their lexicons. What is more, the semantic relatedness between words that children learn from child-directed speech and that serve to link the words within their lexical networks might differ from one family to another. This poses a link between language processing skills and the structure of the language environment. As a first step into shedding some light on this question, the current study examines the language structure experienced by children who are known to process language differently to some degree, i.e., late talkers. Finding evidence of poorly-structured parental speech and related poorly-structured vocabularies is important in that it could motivate future experimental research to confirm the influence that language environment might have on the children's vocabulary structure, which in turn influences how children process words (environment → vocabulary connectivity → language processing).

4.1.3 Differences in the Language Environment

Many studies have investigated what differentiates the parental speech of late talkers from that of typical talkers. Most research conducted on maternal speech style has found no significant differences between the language input received by late talkers and typical talkers in many qualitative and quantitative measures (Paul & Elwood, 1991; Rescorla & Fechnay, 1996). When differences were identified, the authors suspected that they could be a parental adaptation to the child's verbal abilities (pragmatic language interactions: Whitehurst et al., 1988; expansion and extension: Paul and Elwood, 1991; imitation and expansions: Girolametto et al., 1999; responses, expansions and self-directed speech: Vigil et al., 2005). In contrast, degrees in diversity and quantity of the language input have been found to be associated with rates of lexical development. In a prominent study, Hart and Risley (1995) found rapid vocabulary growth in children whose caregivers provided more language input overall, which positively influences vocabulary acquisition (Schwartz & Terrell, 1983). The degree of word diversity in maternal speech was found to positively influence the vocabulary growth of 2-year-old children (Hoff & Naigles, 2002). Recent modelling work suggests that the language outcomes of children with a resolving delay might be more influenced

by the characteristics of their linguistic environment than the language outcomes of children with persisting delay (Thomas & Knowland, 2014). In particular, Thomas and Knowland (2014) measured the richness of the environment as a compound of quantity and quality (diversity of the word types), which leaves unanswered the question of whether either quality or quantity has a greater influence over the child's language outcome.

Whether insufficient semantic richness (operationalized as contextual diversity) of the language environment impedes typical language development is still unknown. What we know is that contextual diversity is lower in language directed at children than when directed at other adults (Hills, 2013). In addition, contextual diversity in the language directed to younger children is lower than the language directed to older children (Hills, 2013). This demonstrates that parents can adapt the contextual diversity of their language to their child's abilities, possibly in an effort to facilitate word learning by increasing contextual consistency, something that has been shown to promote word acquisition (Roy et al., 2015). At the same time, children benefit from the extra complexity in the environment (i.e., diversity). Parents appear to notice this advancement and increase complexity in their speech, thus beginning a virtuous cycle of improvement. This transition from consistency to diversity throughout early childhood requires parents to be sensitive to their child's linguistics needs. This implies that parents that keep their speech too simple (i.e. low in contextual diversity) may delay their child's access to the different semantic nuances of words critical for learning their meaning.

Contextual diversity is not the only semantic-structural feature in language: network attributes in child-directed speech can also be informative. In human language, words are semantically organised forming small-worlds, which together with the presence of a few hub-like nodes shortens the distance between any pair of words in the linguistic system (Ferrer i Cancho & Solé, 2001). We assume most children experience a baseline of these structural characteristics; however, it remains unexplored as to what the optimal structural attributes of child-directed speech are, those which would make children learn faster or influence the type of words to be acquired. As a first step, the current study investigates the association between these environmental features and the rate of lexical development by characterizing the semantic properties of child-directed speech experienced by typical talkers, late talkers, and late bloomers.

4.1.4 Current study

The current study aims to explore the potential implications that individual linguistic environments have on early language development, particularly on the rate of lexical acquisition and language processing. We tracked the expressive (production) and receptive (comprehension) vocabularies of 63 toddlers alongside routine audio samples of the natural language they experienced over a period of six months. Out of this sample, 19 are typical talkers (TT), 22 are late talkers (LT), and 22 are late bloomers (LB). Our main research questions are the following:

Is there an association between the degree of the semantic richness in the linguistic environment of individual children and their rate of lexical development? To explore the relationship between environment and rate of lexical acquisition, we calculated the semantic quality of each type of talker's environment by aggregating all the audio transcriptions into three corpora by type of talker (LT, TT, LB) and then calculated the contextual diversity for each word that the three corpora had in common. Given the important role that contextual diversity has been shown to have on word learning, we hypothesized that the caregiver speech experienced by children with language delay (i.e., LTs and LBs) would be lower in contextual diversity than for TTs. In addition, we also examined the association of contextual diversity in child-directed speech with the child's age and vocabulary size. Although we expect a similar outcome to Hills (2013)—with younger children receiving less contextually diverse input than older children—we also predict based on prior work that parents will adapt their speech to their child's linguistic competence (e.g., Dykstra et al. 2012; Hani, Gonzalez-Barrero & Nadig 2013; Paul & Elwood 1991), such that, for example, older later talkers will still see less contextually diverse language environments than younger typical talkers who know more words.

Are there semantic-structural differences between the speech that late talkers, late bloomers and typical talkers receive at home? We computed and examined three network properties for each corpus (i.e., LT, TT, and LB) that evaluate the general connectivity of the lexical environment. Network analysis in cognitive psychology has been effective in identifying structural differences in language acquisition (Bilson, Yoshida, Tran, Woods, & Hills, 2015; Hills, Maouene, Maouene, Sheya & Smith, 2009a) and lexical processing (Vitevitch, Chan & Goldstein, 2014; Vitevitch, Ercal & Adagarla, 2011; also see Vitevitch, 2019). Our analysis at this point is merely exploratory since no study to date has previously hypothesized the structural attributes of the speech that could be

associated to slow lexical growth. However, as we predict environmental differences based on contextual diversity, we also expect that other related structural features will differ between our corpora.

Do lexical associations learned from unique environments have a significant impact on how children's vocabularies are structured? To be more specific, do all children generally learn the same type of semantic associations between words from their individual environments? If the answer is yes, the network structure of their vocabularies should remain similar, i.e., vocabularies are equally well-connected across talker types. If the answer is no, the differences in the connectivity across groups could theoretically influence the way in which they process language. For example, children with a poorly-connected vocabulary might take longer to navigate through their lexicon in a search for a word. Still, we also consider that late-talking toddlers might acquire a different set of words compared to vocabulary-size-matched typical talkers (Beckage et al., 2011). To explore this hypothesis, we also connected the words within each child's network using the lexical associations extracted from the whole corpus, where all the children's language environments were incorporated.

4.2 Method

4.2.1 Participants

Ethical permission was granted by the Department of Psychology Research Ethics Committee at the University of Warwick to conduct the current research. The families were recruited (with specific emphasis on recruiting late-talkers) through the child-laboratory database of the University of Warwick's baby lab, with study adverts in parental groups on Facebook, UK parental websites and some local nurseries. One hundred fourteen British families participated in the six-month-long study, after excluding 15 families stopped before the study ended. Sixty-three children and their families remained in the study after applying the following exclusion criteria: We excluded families whose children experienced major medical problems, were treated for an ear infection for a prolonged period of time or more than once, had a diagnosed developmental disability, visual or hearing impairment, or were families who spoke more than one language at home. We also excluded those families whose children had a vocabulary size at the end of the study larger than the largest vocabulary size registered for the late talking group (necessary to vocabulary-match our typical group with the two

language-delayed groups). In addition, we only included families that participated for a minimum of three months and audio recorded at least 3 of the five requested topics.

The vocabularies of children were evaluated using a UK adaptation of the MacArthur Child Development Inventory Words & Sentences (W&S CDI, Fenson et al., 1994) created by a group of researchers at the University of Lincoln, UK (Meints & Flecher, 2001). The word "church" was modified to be "church/mosque/synagogue/temple" in order to be more inclusive. The vocabulary checklist is available in the Online Supplementary Materials. To date, there are not any published norms collected from British children older than 25 months old, which prevent studies from including late talkers with large vocabularies. Therefore, to identify late-talking children in our sample, we used Fenson et al.'s (1994) W&S CDI vocabulary norms. In creating British norming data for British infants (12 to 25 months old), Hamilton, Plunkett and Schafer (2000) found that British children showed lower comprehension and production vocabularies than American children. This finding implies that any application of American norms on British children might be inaccurate. Consequently, we acknowledge that some late talkers we identified in our study might not have an actual language delay in a British environment. However, the use of American norms allowed us to determine within our sample which children are at the bottom of the vocabulary spectrum, i.e., which children have the smallest vocabulary relative to their age. Compared to the option of using the UK W&G CDI and their word learning norms for British children aged 8 to 18 months (Alcock, Meints & Rowland, 2020), an additional advantage of using W&S CDI (Fenson et al., 1994) is that we could evaluate the vocabulary of older late talkers, allowing us to examine the development of their vocabularies up to larger sizes.

We identified late talkers in our sample as those whose productive vocabularies are at or below the 20th percentile. We chose the 20th percentile criterion following previous work in semantic networks in late-talking children (Beckage et al. 2011). We assigned two percentiles to each child, one that corresponds to their vocabulary at the beginning of the study, and a second one that corresponds to their vocabulary at the end of the study. Those children who were identified as LT at the beginning of the study and TT at the end of the study were allocated to the late bloomer group. We used the W&S CDI norms to assign a percentile to each child, except for 19 children at the beginning of the study who started younger than 16 months old, in which case the W&G CDI was used (only words in the W&G CDI were considered for computing the

productive vocabulary of these 19 children before assigning their corresponding percentile). One LT was beyond the age of the normative data on the CDI W&S at the end of the study (33.5 months old); his production vocabulary (127 words) was well under the 5th percentile for a 30-month old child, and by extrapolation, we labelled this child as LT. We classified 22 children as late talkers (female=6), 22 children as late bloomers (female=12) and 19 children as typical talkers (female=10). There were no significant differences between the LTs', TTs' and LBs' families in terms of maternal and paternal education, maternal and paternal age, household income per year, number of siblings, attendance to nursery and exposure to Baby signs. For a detailed description of the participating families, see Appendix F

Table 4.1 shows the children's characteristics in each group. We conducted a series of simple linear regression to identify any differences between the groups. Late talker is the oldest group, and differed in age from LBs ($p < .01$, 95% CI [-4.5, -0.83]) and TTs ($p < .001$, 95% CI [-5.7, -1.89]); LBs had a higher average age than TTs, however this difference was not significant ($p > .05$, 95% CI [-0.8, 3.1]); $R^2 = .20$, $F(2, 60) = 8.498$, $p < .001$, $d = 0.18$. Regarding the number of words produced at the beginning of the study, LTs and LBs showed comparable vocabulary sizes ($p > .05$, 95% CI [-0.02, 0.01]; data transformed using a Tukey's Ladder of Powers approach), while TTs presented larger vocabularies than both LTs ($p < .01$, 95% CI [-0.039, -0.009]), and LBs ($p < .01$, 95% CI [-0.04, -0.010]); $R^2 = .16$, $F(2, 60) = 6.962$, $p < .01$, $d = 0.15$. However, at the end of study LBs produced a similar number of words as TTs ($p > .05$, 95% CI [-88.2, 31.0]), and both differed to LTs, who showed the lowest word production (vs TTs: $p < .001$, 95% CI [91.3, 210.5]; vs LBs: $p < .001$, 95% CI [64.9, 179.7]); $R^2 = .31$, $F(2, 60) = 15$, $p < .001$, $d = 0.30$. Likewise, the comprehension vocabularies of LTs and LBs began at similar levels ($p > .05$, 95% CI [-0.62., 0.37], data transformed using a Tukey's Ladder of Powers approach), and different to TTs, who understood more words (vs LTs: $p < .05$, 95% CI [-1.17, -0.14]; vs LBs: $p < .01$, 95% CI [-1.29, -0.26]); $R^2 = .12$, $F(2, 60) = 5.143$, $p < .01$, $d = 0.10$. At the end of study, LBs had a comparable comprehension vocabulary to that of TTs ($p > .05$, 95% CI [-156.8, 40.15]), and both groups differed to LTs (vs TTs: $p < .001$, 95% CI [129.8, 326.7]; vs LBs: $p < .001$, 95% CI [75.1, 264.7]), whose comprehension vocabularies were the smallest; $R^2 = .26$, $F(2, 60) = 11.94$, $p < .001$, $d = 0.25$.

Table 4.1 Description of the Participating Children

	Typical talkers (19)		Late talkers (22)		Late bloomers (22)		Full sample	
	Mean (*SD*)	Range	Mean (*SD*)	Range	Mean (*SD*)	Range	Mean (*SD*)	Range
Age (months)								
Start of the study	17.0 (3.3)	13.0-23.4	20.7 (3.5)	15.6-28.0	18.0 (2.3)	15.1-22.3	18.2 (3.3)	13.0-28.0
End of study	22.1 (3.0)	18.5-28.4	25.9 (3.4)	20.2-33.5	23.7 (2.4)	20.0-27.8	24.0 (3.3)	18.5-33.5
Productive vocabulary size								
Start of the study	58.8 (58.3)	5-223	19.6 (22.2)	2-73	15.2 (10.8)	2-49	29.9 (39.1)	2-223
End of study	269.0 (111.0)	61-383	118.0 (94.2)	21-392	241.0 (80.2)	68-368	206.7 (114.6)	21-392
Receptive vocabulary size								
Start of the study	156.0 (121.0)	5-460	79.9 (61.1)	3-265	71.5 (75.8)	3-358	99.9 (94.1)	3-460
End of study	547.0 (163.0)	298-881	319.0 (131.0)	50-567	489.0 (175.0)	263-903	447.2 (182.8)	50-903
Number of vocabulary updates	26.6 (17.8)	7-68	18.1 (14.1)	4-56	35.9 (25.2)	4-113	26.9 (20.9)	4-113

4.2.2 Procedure

The study had a duration of six months. Parents were instructed to download a specially designed application onto their smartphones or tablets for the study. Parents were then asked at the beginning and throughout the study to mark the words that their child either "says and understands" or just "understands". There were no differences between LTs and TTs in terms of vocabulary updates (see Table 4.1; $p > .05$, 95% CI [-0.01, 0.16]; data transformed using a Tukey's Ladder of Powers approach); the families of LBs updated their vocabularies checklist more often than the families of LTs ($p < .01$, 95% CI [0.04, 0.21]); LBs' and TTs' families updated their children's vocabulary checklists a comparable number of times ($p > .05$, 95% CI [-0.14, 0.03]); $R^2 = .11$, $F(2, 60) = 4.869$, $p < .05$, $d = 0.10$.

Parents were also instructed to record audio of themselves (using the app) while interacting with their children during different daily routines (two audio recordings for mealtime, bedtime, bath time, and nappy/potty time, and four audio recordings of playtime, with one recording every fortnight). All audio files were transcribed by a professional UK-based transcription company. Due to the high cost of transcribing all the audio data from these families, we selected only two audios of playtime per family for transcription. All the other audio files were transcribed. A total of 497 audio files were transcribed. For more details about how data was collected, see the Online Supplementary Materials.

4.2.3 Corpus Cleaning and CDI Words

All utterances produced by children were removed as we are interested in child-directed speech. Punctuation marks were deleted and contracted words were divided and properly corrected (e.g., *you've* was changed to *you have*). We corrected misspellings by running an automated spelling checker and then we removed words with less than two occurrences in the corpus. All words in the corpus were lemmatized and then stemmed. Out of the 676 CDI words, we identified 513 in the corpus after exclusions. We excluded homonyms from the network analysis, e.g., *dry* as an adjective and *dry* as a verb; we also excluded words with the same semantic root, e.g., *sleep* and *sleepy,* since these words resulted as being the same after stemming the corpus (we stemmed all 676 CDI words and then identified those what were not unique in the word set). For those CDI items that included two words, e.g., *can/tin*, we kept the most frequent of the two

in the corpus. The final sample of words for analysis included 95 action words, 318 nouns, 60 adjectives, 21 function words and 19 words related to games/sounds (e.g. baa-baa or peekaboo) and routines (e.g. hi or bye).

4.2.4 Word Co-occurrences

A word's lexical context is constituted by those words that frequently surround it within adult speech. Based on prior work (Hills et al., 2007; Hills, 2015), a window of size five was moved though the corpus to compute the number of times that each word type co-occurred with other words (surface proximity approach; Evert, 2008). Results were stored in a matrix, where the first word in the window indexed the row [i,], and all encountered words (i.e., all other words within the window) indexed the columns [, j]. The resulting weighted matrix was transformed into a binary matrix using a threshold of >0. Words that co-occurred in the corpus are linked by a 1 in this adjacency matrix, representing the semantic relatedness between connected words.

4.2.5 Contextual Diversity and Semantic Networks

In the present study we measure the word's *contextual diversity* as the number of unique word types that appears near the word in question, e.g. within five words (also known as the *lexical environment*, McDonald & Shillcock, 2001). This contrasts with other measures of contextual diversity that consider whole documents or text passages as linguistic contexts (e.g., Adelman, Brown & Quesada, 2006; Hoffman, Ralph & Rogers, 2013). To our knowledge, studies that directly examined the semantic structure of children's lexicons and used child-directed speech as a source for identifying the semantic associates of the words (i.e., contextual diversity when the number of associates is summed) have employed a window-context framework, an approach that we use in the current study (Beckage et al., 2011; Hills et al. 2010; Jiménez & Hills, 2017).

To compute the contextual diversity value for each word and to construct lexical networks, the adjacency matrix described in the last section was used. Depending on the research enquiry, either the whole corpus or a subsample of the corpus was utilised (we describe the details in the next section). The contextual diversity value for each word was calculated by adding the sum of the word's row and the sum of word's column in the matrix.

We also analyzed the network structure of the language environment as well as that of the children's vocabularies by considering words as nodes/vertices and the links between words as a way of representing the semantic relatedness between the words. Undirected networks were built from the adjacency matrix described above, and three structural properties were calculated from them: average degree, local clustering coefficient, and average path length. The *degree* of a node represents the number of ties it has with other nodes. Averaging the degree of all the nodes in the network can give us an idea of the level of cohesion. The *local clustering coefficient* evaluates how well connected the neighbours of a node are among one another. This measure describes not just the connectedness of the network but also the presence of semantic clusters of words in the child's vocabulary. Lastly, the *average path length* measures the average of the shortest path between all pairs of words in a network, providing the degree of its global access. These three network properties are often used in network science to assess the state of connectivity of networks and are also known to differ between early and late talkers (Beckage et al., 2011). We used R and the *igraph* package to compute all network properties (version 1.0.1; Csárdi & Nepusz, 2006).

4.2.6 Corpus Analysis

Each document (transcription) was assigned with the vocabulary size, age, and talker type (LT, TT or LB) of the associated child. We split these files by talker type creating three corpora. The TT group had 144 documents with a total of 82,984 tokens, the LT group had 166 documents with 65,373 tokens, and the LB group had 187 documents with 91,349 tokens. That is, the total number of tokens contained in each talker-related corpus differed significantly, which considerably affects the opportunities of two words to co-occur. Therefore, to control for the size of the corpora, we conducted a population sampling technique. We randomly sampled documents from each corpus until the total number of words accumulated reached a threshold. The threshold used was half the total number of words in the smallest corpus (i.e., LT's corpus, threshold= 32,686; we obtained similar results when the threshold used was one-third of the total number of words in the smallest corpus). To make sure that the three randomly sampled corpora had exactly 32,686 tokens, we trimmed the excess words from the end of the last document/transcription sampled. Then, we created three adjacency matrices with the words that the three sampled corpora had in common. The contextual diversity for each word was calculated for each matrix and

then averaged for the sample. Network statistics were also produced for each sample. In addition, the age and vocabulary size associated with the sampled documents were averaged for each group. We repeated these steps 1,000 times for each corpus. This means that for each corresponding talker-type corpus, we produced and recorded 1,000 average contextual diversity values, 1,000 average degree values, 1,000 clustering coefficients, 1,000 average path length values, 1,000 average age values, and 1,000 average vocabulary size values.

For the statistical analysis, we conducted a standard stepwise regression. The *lm* function in R was used (R Core Team, 2019). Heteroscedasticity was detected, which could cause the standard error to be biased for model comparisons. Therefore, we performed a heteroscedasticity robust F-test to compute robust standard errors. The Wald test was selected, which relaxes the assumption of errors being independent and identically distributed. We used the *waldtest* function in the R package *lmtest* (Zeileis & Hothorn, 2002).

4.2.7 Network analysis

To examine whether potential differences in the network properties of the expressive (production) and receptive (comprehension) vocabularies of TTs, LTs and LBs are due to the acquisition of different sets of words, we constructed the children's networks using the same adjacency matrix computed from the whole corpus. In contrast, to examine the question of whether differences in the network properties of children's vocabularies are potentially caused by differences in their individual environments, we also constructed the children's networks using unique adjacency matrices computed from their corresponding corpora, i.e., LT, TT, and LB. However, as we pointed out above, the three corpora differed in size which would likely impact results. To solve this issue, we resorted to population sampling again. In fact, we used the same matrices generated for our corpus analysis to construct the children's networks. This means that for each child we computed 1,000 values for each of the three network properties considered in this study (average degree, clustering coefficient and average path length), and then averaged them for each child.

For the statistical analysis, we utilised generalized additive models. The *gam* function in the *mgcv* package in R was used (see Wood 2001). Generalized additive models were selected over simpler statistical analysis as its smooth functions allowed us

to relax assumptions that were found to be violated. First, heteroscedasticity was detected: as vocabulary size increases, the differences between the observations and the regression line became larger. Second, there were differences in the vocabulary size of LTs, TTs and LBs (see Table 4.1), and vocabulary size was highly correlated to our independent variable. With GAMs we dealt with this issue by adding vocabulary size as a smoothing term in our GAM to control for it. Third, simpler regressions showed a poor fit to the data due to a high variance at early stages of vocabulary development, where GAMs offer a better performance thanks to local fits. To build our model, we added predictors in a hierarchical fashion as fixed effect terms, and then we identified the best model comparing their BICs. Vocabulary size was entered as a smooth term and random smooths were introduced by participants to take into account the repeated measures.

4.3 Results

4.3.1 Semantic Structure of Child-Directed Speech Across Talker Types

We examined whether parental speech varies in semantic structure depending on whether it is directed to LTs, LBs or TTs. Table 4.2 displays the results from the best regression models selected after stepwise comparisons (BICs compared). The independent variables tested were talker type, vocabulary size and age. All winning models included talker type as a predictor, indicating that all semantic structural measures differed across the talker's environments. Standardized regression coefficients for all four models are plotted in Figure 4.1. Given that we are considering all words in the corpus (not just words in the CDI) to calculate networks statistics of the environment, contextual diversity and average degree become practically the same measure. Contextual diversity and average degree generally increase as children age, consistent with Hills' results (2012). The language environment of LTs are lower in contextual diversity and average degree than in TTs' and LBs' environments, with the TT and LB environments showing similar quantities. Clustering coefficient decreases as children age; it also generally decreases as vocabularies grow following a short rise at the

beginning[1]. The language environment of TTs shows the lowest clustering coefficient of the three groups, and LBs the highest. All groups significantly differed from each other. Average path length in the language environment generally increases as children produce more words. The language received by LTs shows the highest average path length values, followed by TTs, and LBs with the lowest average path length in the environment. In sum, the semantic structure of parental speech differs across the type of talkers to which it is directed to.

Table 4.3 shows the Pearson correlations between all network properties. As expected, average degree and contextual diversity are strongly and positively correlated. These two properties are highly negatively correlated with average path length, and weakly negatively correlated with clustering coefficient. Clustering coefficient is weakly negatively correlated with averaged path length.

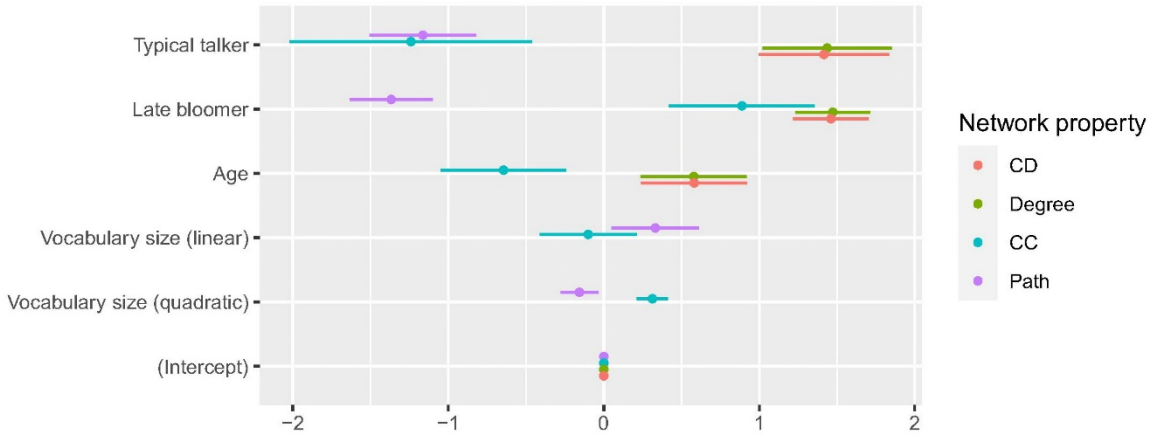
[1] Due to potential collinearity between age and vocabulary, we checked the coefficients of separate models, each one predicting clustering coefficient from either age or vocabulary size as well as talker type. The signs of both coefficients coincide with those reported in Table 4.2.

Table 4.2 Regression Analysis Predicting Contextual Diversity and Network Properties in Child-Directed Speech

Dependent Variable	Independent Variables	*b*	*CI*	*SE*	*p*	*F (df)*	*Adj.R²*
Contextual Diversity	Age	0.29	[0.20, 0.46]	0.09	< .001	252.4 (3,2996)	0.20
	Talker type	TT (vs LT): 1.42	[0.99, 1.83]	0.21	< .001		
		LB (vs LT): 1.46	[1,22, 1,70]	0.12	< .001		
		TT (vs LB): -0.05	[-0.24, 0.15]	0.10	>.05		
Average Degree	Age	0.18	[0.07, 0.29]	0.05	< .001	262.3 (3,2996)	0.20
	Talker type	TT (vs LT): 1.44	[1.02, 1.85]	0.21	< .001		
		LB (vs LT): 1.47	[1.23, 1.71]	0.12	< .001		
		TT (vs LB): -0.04	[-0.23, 0.16]	0.10	>.05		
Clustering Coefficient	Age	-0.20	[-0.32, -0.07]	0.06	< .01	464.6 (5,2994)	0.44
	Vocabulary size	Linear: -2.75	[-11.29, 5.79]	4.36	> .05		
		Quadratic: 8.56	[5.83, 11.27]	1.39	< .001		
	Talker type	TT (vs LT): -1.24	[-2.02, -0.46]	0.40	< .01		
		LB (vs LT): 0.89	[0.42, 1.36]	0.24	< .001		
		TT (vs LB): -2.13	[-2.46, -1.79]	0.10	< .001		
Average Path Length	Vocabulary Size	Linear: 9.08	[1.37, 16.79]	3.93	<.05	174.1 (4, 2995)	0.19
		Quadratic: -4.28	[-7.55, -1.01]	1.67			
	Talker type	TT (vs LT): -1.16	[-1.51, -0.82]	0.18	< .001		
		LB (vs LT): -1.37	[-1.63, -1.10]	0.14	< .001		
		TT (vs LB): 0.20	[0.06, 0.35]	0.08	<.01		

Note. Models with the lowest BIC are displayed. Tukey's Ladder of Powers was conducted to transform values in the dependent variable to comply with regression's assumptions. Values are standardised.

Figure 4.1 Standardised Regression Coefficients of Best Models Predicting Contextual Diversity and other Network Properties



Note. CD= contextual diversity. Degree= average degree. CC= clustering coefficient. Path= average path length. The referent group is late talker. Details of the best models can be seen in Table 4.2.

Table 4.3 Pearson Correlations Among Network Properties in Child-directed Speech

	Variable	1	2	3	*4*
1	Average degree	1.00			
2	Contextual diversity	0.99	1.00		
3	Clustering coefficient	-0.12	-0.12	1.00	
4	Average path length	-0.65	-0.66	-0.23	1.00

4.3.2 Network Properties of Vocabularies

We also asked whether the network properties of LT vocabularies differed from that of TT and LB vocabularies. We explored this question by conducting separate GAMs (see details in Method). Data points with less than four words were removed to reduce variance. Figure 4.2 shows the growth of average degree (top), clustering coefficient (middle), and average path length (bottom) of the expressive lexical networks of LTs, TTs and LBs. On the left-hand side, the same adjacency matrix (produced from the whole corpus) was used to build the children's networks, which we refer to as 'same environment'. On the right-hand side, properties of the same vocabularies are displayed but networks were built using their corresponding talker-type-related adjacency matrix, which we refer to as 'different environment'. Figure 4.3 shows the same network properties but for the children's receptive vocabularies. In general, for both expressive and receptive vocabularies, the way in which all three network measures develop with increasing vocabulary size is consistent with previous work (Beckage et al., 2011, Bilson et al., 2005): average degree and average path length increases, whereas clustering coefficient decreases as vocabularies grow.

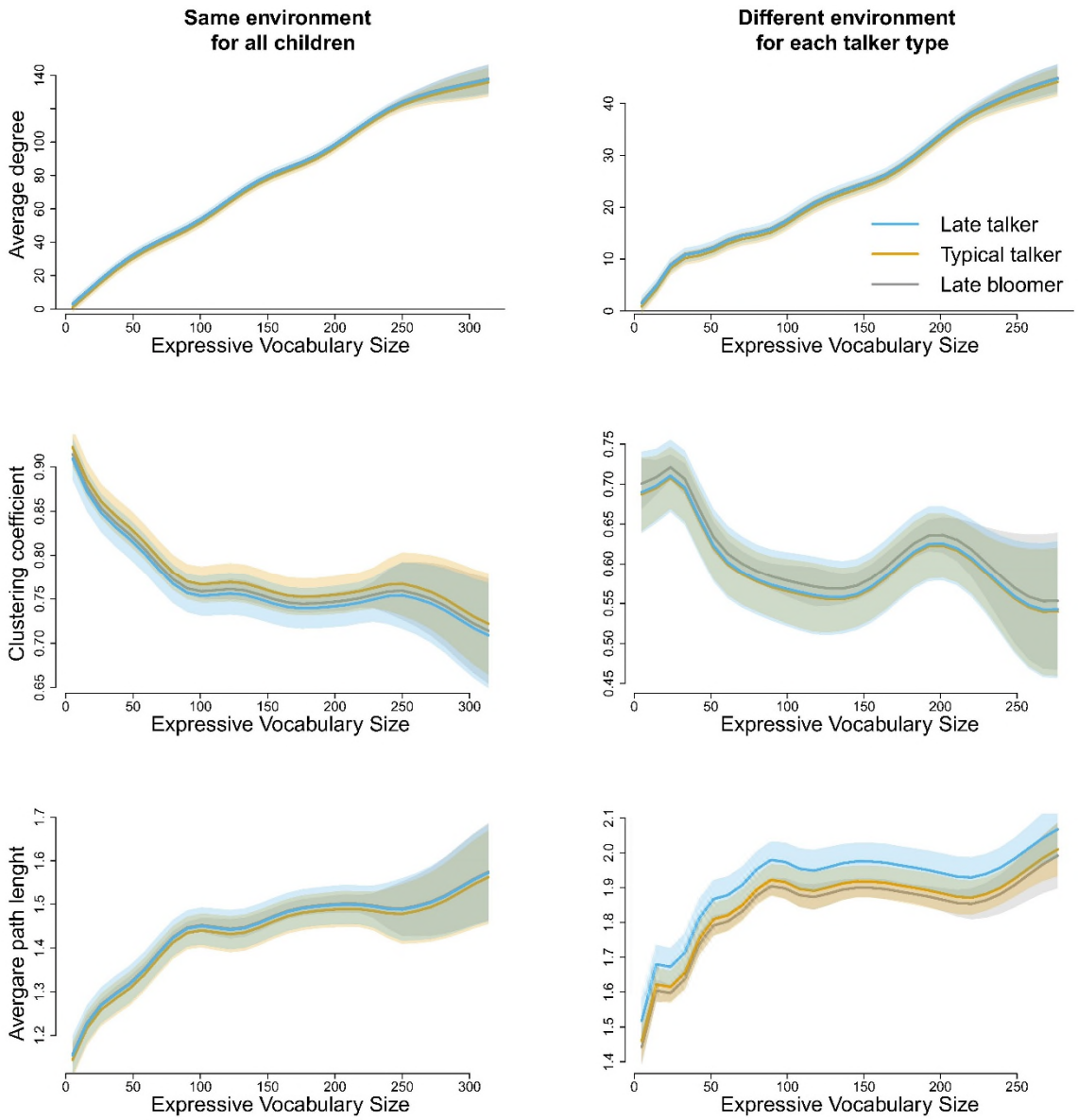
Table 4.4 shows the results of GAMs predicting each network property. Although the models include talker type as a predictor, only a few of these models showed a significant improvement in the predictive power over a simpler version of the model without talker type as a predictor, indicated in Table 4.4 with an asterisk. When considering the same language environment for all children (results shown at the top section of Table 4.4), network properties of children's expressive vocabularies (left-hand-side of Table 4.4) were found to be similar across talker types. In contrast, the network properties of children's receptive vocabularies differed across talker types in two network properties: LBs exhibited larger clustering coefficients than TTs and LTs, who did not differ between them; and LTs presented larger average path distances than LBs and TTs, with LBs and TTs showing similar values of path length. These two models predicting clustering coefficient and average path length significantly increased the predictive power of a simpler model after adding talker type as a predictor. No differences across talker types were found for average degree.

When considering different language environments for each talker type (results showed at the bottom section of Table 4.4), the expressive vocabularies of LTs showed higher average path length than the vocabularies of LBs and TTs; LBs and TTs showed

equivalent average path length values. This model showed significantly higher predictive power than a simpler model without talker type as a predictor. Average degree and clustering coefficient were found to be similar across talker types for their expressive vocabularies. With regard to receptive vocabularies, LBs showed higher average degree and higher clustering coefficient than LTs, however, none of these two models improved the predictive power of simpler models. Average path length was identical across talker types.

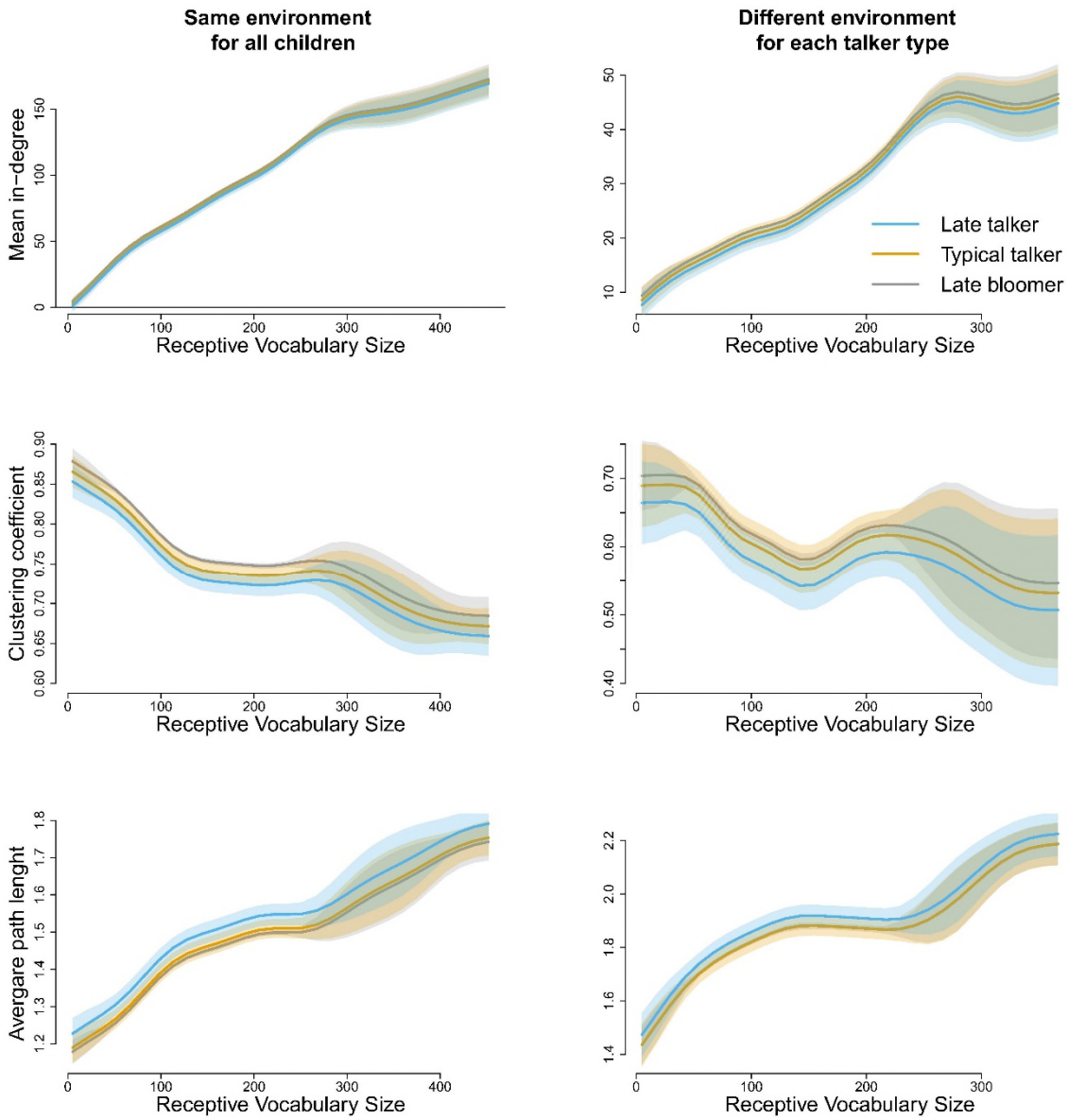
In sum, LTs, TTs and LBs understand a different set of words that lead to differences in clustering coefficient and average path length of their receptive networks, but not in their expressive networks. In contrast, the learning of unique semantic associations between words from their individual talker-related environments lead to differences in average path length in their expressive vocabularies, but not in their receptive vocabularies.

Figure 4.2 Network Properties of the Expressive Vocabularies of Typical Talkers, Late Talkers, and Late Bloomers across Different Vocabulary Sizes.



Note. Networks on the left-hand side were built using the same adjacency matrix generated from the whole corpus. Networks on the right-hand side were built using the unique adjacency matrices generated from either the LT corpus (for LT vocabularies), TT corpus (for TT vocabularies), and LB corpus (for LB vocabularies). Smoothed plots are based on the GAM models predictions. Shadows around the curves represent 95% confidence intervals.

Figure 4.3 Network Properties of the Receptive Vocabularies of Typical Talkers, Late Talkers, and Late Bloomers across Different Expressive Vocabulary Sizes.



Note. Networks on the left-hand side were built using the same adjacency matrix generated from the whole corpus. Networks on the right-hand side were built using the unique adjacency matrices generated from either the LT corpus (for LT vocabularies), TT corpus (for TT vocabularies), and LB corpus (for LB vocabularies). Smooth plots are based on the GAM models predictions. Shadows around the curves represent 95% confidence intervals.

Table 4.4 Regression Analysis Predicting Network Properties in Children's Vocabularies: Semantic Relatedness Calculated from Whole Corpus Versus from Distinct Talker-type-related Corpora

		Expressive vocabulary						**Receptive vocabulary**					
Environment	**Dependent Variable**	***b***	***CI***	***SE***	***p***	***F (df)***	***Adj.R²***	***b***	***CI***	***SE***	***p***	***F (df)***	***Adj.R²***
Same Environment	**Degree**	TT (vs LT): -1.64	[-4.88, 1.59]	1.61	> .05	1.213	0.99	TT (vs LT): 1.55	[-2.27, 5.36]	1.90	> .05	1.88	1
		LB (vs LT): 0.53	[-2.54, 3.59]	1.52	> .05	(2,1365)		LB (vs LT): 3.37	[-1.18, 6.91]	1.76	> .05	(2,1370)	
		TT (vs LB): -2.17	[-5.04, 0.70]	1.42	> .05			TT (vs LB): -1.82	[-5.28, 1.63]	1.70	> .05		
	CC	TT (vs LT): 0.013	[-0.01, 0.04]	1.18	> .05	0.74	0.92	TT (vs LT): 0.012	[-0.002, 0.027]	0.007	> .05	7.61*	0.97
		LB (vs LT): 0.005	[-0.02- 0.03]	0.48	> .05	(2,1424)		LB (vs LT): 0.025	[0.012, 0.038]	0.007	< .001	(2,1398)	
		TT (vs LB): 0.007	[-0.012, 0.027]	0.009	> .05			TT (vs LB): -0.013	[-0.025, -0.001]	0.006	< .05		
	Path	TT (vs LT): -0.010	[-0.06, 0.04]	0.021	> .05	0.25	0.94	TT (vs LT): -0.038	[-0.068, -0.007]	0.015	< .05	6.25*	0.98
		LB (vs LT): 0.002	[-0.04, 0.04]	0.020	> .05	(2,1422)		LB (vs LT): -0.049	[-0.077, -0.021]	0.014	< .001	(2,1398)	
		TT (vs LB): -0.012	[-0.051, 0.027]	0.018	> .05			TT (vs LB): 0.012	[-0.014, 0.036]	0.012	> .05		
Different Environment	**Degree**	TT (vs LT): 0.19	[-2.04, 0.90]	0.68	> .05	0.76	0.99	TT (vs LT): 0.86	[-0.97, 2.70]	0.91	> .05	2.07	0.99
		LB (vs LT): -0.57	[-1.19, 1.58]	0.73	> .05	(2,1336)		LB (vs LT): 1.71	[0.011, 3.41]	0.84	< .05	(2,1360)	
		TT (vs LB): -0.77	[-0.50, 2.04]	0.63	> .05			TT (vs LB): -0.85	[-0.84, 2.54]	1.01	> .05		
	CC	TT (vs LT): -0.003	[-0.048, 0.042]	0.018	> .05	0.31	0.85	TT (vs LT): 0.025	[-0.013, 0.064]	0.019	> .05	2.52	0.92
		LB (vs LT): 0.010	[-0.033, 0.054]	0.021	> .05	(2,1372)		LB (vs LT): 0.039	[0.004, 0.075]	0.018	< .05	(2,1389)	
		TT (vs LB): -0.013	[-0.024, 0.050]	0.018	> .05			TT (vs LB): -0.014	[-0.050, 0.021]	0.017	> .05		
	Path	TT (vs LT): -0.58	[-0.102, -0.013]	0.022	< .01	5.25*	0.89	TT (vs LT): -0.037	[-0.083, 0.009]	0.023	> .05	1.93	0.93
		LB (vs LT): -0.076	[-0.123, -0.027]	0.024	< .01	(2,1380)		LB (vs LT): -0.038	[-0.078, 0.002]	0.019	> .05	(2,1382)	
		TT (vs LB): 0.018	[-0.057, 0.021]	0.019	> .05			TT (vs LB): 0.001	[-0.037, 0.035]	0.018	> .05		

Note. * signifies the model significantly improved their predictability after the variable 'type of talker' was added. BICs compared. Adding an interaction between talker type and vocabulary size did not significantly improve the prediction power of any model. CD= contextual diversity. Degree= average degree. CC= clustering coefficient. Path= average path length.

4.4 Discussion

The present study examined the relation between the structural quality of the language environment and the rate of lexical acquisition in children. Crucially, the structural quality of the child-directed speech experienced by LTs, LBs and TTs is different. The best language outcomes, i.e., TTs, are associated with a language environment whose network structure is higher in contextual diversity and connectivity (degree), has lower semantic clustering, and lower distance between words when compared to the LT and LB outcome groups. We predicted and found that the semantic associations learned in a LT environment led to LTs' productive networks having larger distances between words (higher path length), which might have potential consequences in how LTs process language. Furthermore, LTs seemed to understand a different set of words that affect the connectivity of their receptive networks, with generally lower semantic clusters (than LBs) and higher distance between words (than TTs and LBs). Taken together, these results suggest that toddlers with different patterns of language acquisition also experience different language learning environments, which is not only associated with the rate of acquisition but also the order of acquisition.

Parental speech directed to late bloomers resembles that of TTs, with the exception of the level of semantic clustering, which was the highest across environment groups. If the only factor influencing the word learning rate were the quality of the environment, our results would suggest that what is delaying LBs is the level of semantic clustering within their caregiver's speech. The environment associated with the lowest rates of lexical acquisition (i.e., LTs) presented other additional structural characteristics that could be related to their language delay: lower contextual diversity and degree, and higher path length. The high correlation found between contextual diversity (as well as average degree) and average path length suggest that parents that increase contextual diversity in their speech will potentially also reduce the distance between words in the network structure of their speech. In contrast, contextual diversity and clustering coefficient are weakly correlated in our study, indicating that these two semantic features could vary semi-independently within parental speech. Therefore, in our discussion below, we focus on understanding what these two structural features (contextual diversity and clustering coefficient) look like in real language and how they are related to word learning.

Let us consider the example of the word 'dog'. Parents might increase the contextual diversity of 'dog' by using the word in many different and appropriate contexts, for example, when they see a dog in the park, when they see dog food in the supermarket or when they play with toys. These bring opportunities to associate the word 'dog' with many distinct words from

other semantic categories (e.g., 'leash', 'food' or 'doll'), which we call 'different-category situations'. In contrast, other parents might only produce 'dog' in contexts in which the word appears with other words very close in semantic meaning, e.g., only when playing with animal toys or reading animal books, which we call 'same-category situations'. These two different language experiences lead to different semantic structures in language: for different-category situations, the contextual diversity of 'dog' is higher than in same-category situations, whereas in same-category situations the clustering coefficient (i.e., how well-connected are the word's associates between each other) is higher.

Increasing contextual diversity in speech facilitates word learning (Hills et al., 2010; Rosa, Tapia & Perea 2007; Pagán & Nation, 2018) which can be achieved by increasing different-category situations. Nonetheless, we found contextual diversity of the speech directed to younger children to be lower than that directed to older children, confirming findings by Hills (2013). By lowering the contextual diversity in their speech at earlier stages of lexical development parents are making their speech more consistent and predictable, which some suggest facilitates early word learning (e.g., Roy et al., 2015). In contrast to contextual diversity, clustering coefficient declines in parental speech as their children age. In other words, parents reduce same-category situations as their children grow while increasing different-category situations. The benefits of increasing different-category situations/contextual diversity for word learning are straightforward. However, it is not entirely clear how decreasing same-category situations/semantic clustering could facilitate word learning in young children. We argue that by reducing same-category situations while maintaining high contextual diversity parents intensify the contrast between referents as words present different semantic characteristics between each other. We know that words acquired earlier by young children are semantically more distinct between each other than words acquired later, with semantic relatedness measured as shared visual and surface features (Engelthaler & Hills, 2017). In our example above, the acquisition of the word 'dog' might be easier to learn when co-occurring with less semantically-similar words, like 'food', than when presented with other more semantically-similar words like 'giraffe', which could explain why language lower in local semantic clustering might facilitate early word learning. Nevertheless, there is a chance that the increment of semantic clustering in speech that we observed was a parental adaptation in view of the child's language delay. Further research is needed to evaluate this hypothesis.

So far, we have discussed how structural differences in the language input might promote lexical acquisition. Now we turn to how these differences might influence lexical processing in children who learned semantic relatedness between words from their respective talker-related

environments. Our results showed that LT environments might potentially make LTs' expressive lexical network have a significantly larger distance between words, i.e., higher average path length. In network science, this translates into a less efficient distribution of information throughout the system at a global level (Barabási 2016). For example, after the activation of a target word 'dog', distant word associates would receive a weaker signal in a LT network than in a TT network, which might indicate a less rich semantic representation of the activated word. This fits with previous findings that confirm a link between poor linguistic skills and semantic weakness: children with developmental language disorder (DLD) show poorer semantic knowledge about known words compared to age-matched peers (measured by analyzing the drawings and verbal definitions of the items; McGregor, Newman & Reilly, 2002); also poor reading comprehenders show difficulties recalling abstract words, suggesting deficits in their semantic representation (Nation et al., 1999).

We found that the type of words that LTs understand make their receptive networks have lower semantic clustering and higher path length than TTs and LBs. This finding is appealing given that we used the same adjacency matrix to build their lexical networks (i.e., same environment). What is more, these differences reflect the same differences identified in the LT language environments. This coincidence suggests that the semantic structure of the learning environment might influence the type of words that children acquire, at least the words they come to understand. However, all conclusions based on our receptive vocabulary data should be taken cautiously. Determining what a child understands can be difficult, as suggested by a recent study that showed that parental reports do not entirely match with the experimental data testing word comprehension in young children (Moore et al., 2019). Therefore, more fine-grained data (e.g., from eye-tracking) should be used in future research.

There are several potential limitations to this work. Due to the longitudinal nature of our study, we were able to exclude two potential cases of ASD, as parents kindly informed us about the results from professional checks that their children underwent; however, some additional children could turn out to be diagnosed with DLD in the future. Also, the vocabulary norms utilised to identify the children's percentiles comes from an American English population, which means that some LTs identified here would not be identified with a language delay in a British environment (Hamilton, Plunkett, and Schafer, 2000). Future research should use British norms for toddlers up to 30 months when they become available. Furthermore, there were not enough transcriptions per family to conduct a more fine-grained analysis, and instead, we aggregated contexts per type of talker. More abundant audio data per family could have been more informative about the exact lexical associations learned by each particular child in our sample.

Finally, perhaps the most important limitation is that our study can only conclude an association between the quality of the language environment with the rate of lexical acquisition and not a causal relationship. Although evidence for associations is important as a first step into investigating causation, the next steps would be to experimentally prove the potential effect of less well-connected language environments on early word acquisition as well as on language processing.

In conclusion, the present work shows children with different patterns of language delay experience language environments with different structural properties that correspond with those delays. Moreover, we also find that LBs appear to experience environments that are more similar to TTs' environments. These results are consistent with either an influence of the learning environment or an influence of individual differences in children's language learning abilities which are then adapted to by parents. In either case, this work establishes for the first time using large language corpora in a longitudinal study that language environments differ between talker types and thereby establishes a foundation for future work aimed at disentangling the potential influences between language learners and their environments.

Chapter 5: Untangling the Riddle of Contextual Diversity and Contextual Distinctiveness in Early Word Learning

Language structure matters in early word learning. However, how it matters is something of a riddle. Words with higher contextual diversity have been found to be learned earlier. However, paradoxically, words that appear in more distinct contexts are also easier to learn. To resolve this issue, natural language around daily routines was collected from 63 individual British-English speaking families alongside the vocabularies of their early learners, who were aged 13 to 28 months old. Using hierarchical regressions, we compared contextual diversity and distinctiveness measured at both word (lexical) and routine (document) levels. The best predictors were frequency, lexical-contextual diversity, and document-contextual distinctiveness. How the scale of these contextual measures are associated with word-referent mapping and semantic enrichment is discussed.

5.1 Introduction

It is well-known that young children exploit the structural properties of child-directed speech for early word acquisition (e.g., Mintz, Newport, & Bever, 2002; Weisleder & Waxman, 2010; Mccauley, Monaghan, & Christiansen, 2015). However, some of the evidence around this observation presents us with a riddle. One set of observations suggests that words found in more diverse language environments may be learned earlier (Hills, Maouene, Maouene, Riordan, & Smith, 2010; Hills, 2013; Fourtassi, Bian, & Frank, 2019; Saffran, Newport, & Aslin, 1996). However, paradoxically, another set of observations suggest that words found in more distinct language environments are learned earlier (Roy et al., 2015; Hills, 2013; Waxman & Klibanoff, 2000). How these two potentially opposite but related structural properties of language interact to influence word acquisition is not well understood. In the current study, we investigate this question by comparing the impact of these two contextual

features (diversity and distinctiveness) on early word learning in a sample of children for which we have both vocabularies and the language environment they experienced. This approach also allows us to look more closely at alternative methods for defining language context, which turns out to be key to unravelling the riddle.

There are two main approaches to defining language context in the literature. The first approach considers the *lexical environment* near a word, and considers the diversity of other words that co-occur in the proximal context (e.g., within 10 words) of that word (McDonald & Shillcock, 2001). Measures of contextual diversity of the lexical environment have been shown to be an excellent predictor of a word's age of acquisition (AoA)—outperforming frequency—with the earliest words produced by children being higher in contextual diversity than later learned words (Hills et al., 2010; Sailor, 2013; Bilson et al., 2015). The second approach involves larger contexts, usually considered at the level of the *document environment*, and measures the diversity or distinctiveness of a word by counting the number of different documents in which a word appears. One of the earliest studies using this approach found that contextual diversity measured at the document level predicted performance in word-naming and lexical-decision tasks in adult readers better than word frequency (WF) (Adelman, Brown, & Quesada, 2006). A later study redefined this measure to estimate how ambiguous a word is by comparing the degree of similarity of the contexts in which the word appears. This measure was termed *semantic diversity* by its developers (Hoffman, Ralph, & Rogers, 2013). Similar to contextual diversity, semantic diversity was found to influence word processing. For example, children answer faster to words with high semantic diversity in lexical decision tasks (Hsiao & Nation, 2018), read them aloud more accurately (Hsiao & Nation, 2018), and process them more easily based on eye-movements patterns (Pagán, Bird, Hsiao, & Nation, 2020). In our terminology, we consider semantic diversity to be a document-level measure of contextual diversity, with the goal of isolating words that appear across more unique documents.

The countermeasure to the contextual diversity described above is *contextual distinctiveness* (McDonald & Shillcock, 2001). Contextual distinctiveness stands in opposition to the concept of contextual diversity: whereas contextual diversity relates to the number of different contexts in which the word appears, contextual distinctiveness relates to the uniqueness of the contexts in which the word appears. Consistency across contexts has been suggested to aid word learning (Roy et al., 2015; Brown, 2008;

Waxman & Klibanoff, 2000). Despite the negative correlation of contextual diversity and distinctiveness (Hoffman, Ralph, & Rogers, 2013), contextual distinctiveness has also been shown to contribute to the processing of language. Words with high contextual distinctiveness have slower reaction times in lexical decision tasks (McDonald & Shillcock, 2001). Contextual distinctiveness has been found to predict the age of the first production, not only when contextual distinctiveness was measured from linguistic contexts but also when measured from spatial contexts (where in the house the word was spoken) and temporal contexts (what time of day the word was spoken) (Roy et al., 2015). Roy et al.'s investigation (2015) on contextual distinctiveness was motivated by previous findings that showed that the predictability of linguistic patterns in language (named "transactional formats" by Bruner, 1985) provides a stable framework for children that promotes the acquisition of word learning.

The facilitation effects that diversity and distinctiveness have in early word acquisition are well confirmed. However, the antagonistic character of the two contextual measures lead to the inevitable question: how can they both predict age of acquisition? One might wonder whether computing contextual distinctiveness by considering just the immediate words in the lexical environment, would yield similar or different results to those computed at the document level. Similarly, which type of linguistic context (lexical versus document environment) would be the best for computing contextual diversity to predict AoA. The study of these methodological differences might be valuable as they might show signals of the type of learning mechanism being used by the word learner.

An additional challenge to addressing these questions is finding the natural language corpus on which to base these measures. Many studies examining the structure of large, meaningful linguistic contexts (at the document level) have used text documents such as news articles or textbooks (e.g., Adelman, Brown, & Quesada, 2006; Rosa, Tapia, & Perea, 2017; Jones, Johns, & Recchia, 2012; Hoffman, Ralph, & Rogers, 2013; Hsiao & Nation, 2018). In contrast, work like Roy et al.'s utilised a more naturalistic approach to detect the typical activities in the child's natural language environment (2015). Roy et al. split a large corpus of natural speech directed to a single child into 10-minute-long documents, with the assumption that this allowed them to measure the words' contextual distinctiveness. Using this approach, Roy et al. successfully predicted AoA from this measure of contextual distinctiveness. In the present study, we push this one step further by examining caregiver speech around a set

of independent activities, directed to a specific set of children for which we also have longitudinal vocabulary growth.

Theoretically, the main difference between the lexical and document environment is one of size. For the purpose of mapping novel words to their corresponding referents by using a mutual exclusivity strategy, large contexts, such as the description of an activity, reflect what is in the child's 'scene'. In contrast, for the purpose of learning new semantic nuances of known words, such an ample context might be too large; instead, restricted context might be more informative. Consistent with this, Hills et al. (2010) examined the window size for computing contextual diversity that best predicted order of acquisition for words and found that sizes differed by word class, indicating that some words may benefit from document-size windows whereas others appear to be governed by more narrow windowed associations. For example, consider a child hearing the following sentence: "Look, the boy is now riding a spotty horse". In this narrow context, the child might learn new information about the words 'riding' and 'spotty'. But the fact that the sentence is in the broader context of a book about vehicles may be less informative with respect to these two words. Yet it may be useful for learning about functional features of the word 'horse'. For this reason, we argue that when analysing the linguistic context, the size of the context should matter in informing us of a particular learning purpose (word-referent mapping or semantic enrichment of known words). Critically, by determining the best context size to measure contextual diversity and contextual distinctiveness, one could shed some light on the role of these opposing contextual properties in word learning.

5.1.1 Current study

Focusing on the general influence of contextual information on word learning, we calculated contextual diversity and contextual distinctiveness for each word in a large corpus of child-directed speech from 63 families. To explore the best way to measure these two contextual properties, we calculated each of them by considering two types of contexts: the lexical environment (word's co-occurrence with immediate words), and the document environment (word's occurrence across routines or situations). Previous work by Hills et al.' (2010) sought to answer a similar question by trying different window-sizes (up to 100 words). We differ from this work in that we explore larger, meaningful and comprehensive contexts (young children's routines). In addition, we computed the word acquisition order (a measure correlated with AoA) from the productive vocabulary

of the children associated with these families. With this information, we conducted a series of regression models to reveal the type and degree of contribution that contextual diversity and contextual distinctiveness at the level of the document and lexical environment have on word acquisition order.

We hypothesise that the order in which words are acquired will be associated with the degree of contextual diversity of the word, in that high contextual diversity words are learned earlier, as previously found by Hills et al. (2010). Similarly, we hypothesize that earlier words will have a higher degree of contextual distinctiveness, as was previously shown by Roy et al. (2015). However, due to the opposed nature of the two measures, we predict that only one of these two semantic measures will show the influence in word learning described above.

5.2 Method

5.2.1 Sample

The transcriptions of a total of 495 audio files were used in the current study. These audio files were recorded by parents of 63 British-English-speaking toddlers as part of a 6-months longitudinal study. These families were selected out of a pool of 114 families based on a set of exclusionary criteria. The 114 British families were contacted through the child-laboratory database of the University of Warwick's baby lab, digital adverts in UK-based websites for parents, Facebook parental groups, and local nurseries. The following children and families were excluded from this initial pool: children that experienced major medical problems, endured ear infections recurrently, were diagnosed with a developmental disability, visual or hearing impairment, or lived in a multilingual home. After this filter, children with a vocabulary size larger than 392 words were excluded to vocabulary-match the groups analysed in the original study for which the data was intentionally collected.

Parents were instructed to record routine-related audios at home utilising a smartphone application specially devised for the study. All audio files were professionally transcribed and contained the following topics: mealtime, bedtime, bath time, nappy or potty time, and playtime. Along with the audio recordings, the parents were asked to fill a vocabulary checklist during the study. Each time that the child produced a new word for the first time, they had to mark it on the checklist. The

vocabulary checklist was a British adaptation of the MacArthur Child Development Inventory Words & Sentences (W&S CDI, Fenson et al., 1994), which was constructed by researchers at the University of Lincoln, UK (Meints & Flecher, 2001). This checklist was incorporated in the smartphone application for the study to facilitate the data collection. Families covered between 3 and 5 topics for the audio recordings and participated an average of 164.4 days (5.3 months), *SD* = 20.2.

The children associated with these families were aged 13 to 28 months old when they started the study (*M*=18.2, *SD*=3.3), with productive vocabulary sizes ranging between 2 and 223 words at the beginning of the study (*M*=29.9, *SD*=39.1), and between 21 and 392 words at the end of the study (*M*=206, *SD*=114.6). Most parents (88.9%) in the sample reported average or higher income (= or > £20,000, with 4.8% preferred not to answer), and most mothers reported attaining a high education level (82.5%).

5.2.2 Corpus Processing and CDI words

We removed all utterances produced by children in the corpus. A spelling checker was used and misspelt words were deleted along with words with frequency of 1. All words in the corpus were then lemmatised and stemmed. A total of 455 words from the CDI were in the corpus, of which 300 were nouns, 95 verbs, and 60 adjectives. Though here we are primarily interested in word types that have a strong semantic content—and thus we analysed nouns, which have referential meanings, and verbs and adjectives, that have relational meanings (Gentner, 1978)—the results we present here also hold when all words are taken together.

5.2.3 Word features: Contextual Diversity, Frequency and Contextual Distinctiveness

The utterances of the adults in the sample were processed to compute the contextual diversity, contextual distinctiveness, and word frequency for each word. Word frequency was computed by counting the number of times that each word type appears in the respective corpus.

5.2.3.1 Lexical Environment

The words' contextual diversity was calculated using co-occurrence statistics derived from the corpus. An empty matrix was populated by moving a window of size

10 word-by-word through the corpus. Following Hills et al. (2010), we found that a window size of 10 was best at predicting word acquisition order. The first cell of the window indexed the row [i,], and any word within the window indexed the column [, j]. A value of 1 was added to the existing value at that position in the matrix [i, j]. The matrix was converted into a binary matrix, where any value greater than 0 was replaced with 1. A smaller matrix was created with the 455 CDI words we consider here for analysis, and the sum of the row and the sum of the column were added to calculate the contextual diversity value for each word.

In addition to contextual diversity, we computed contextual distinctiveness within the window-context framework (i.e., lexical context). For this purpose, we calculated the entropy of each word from the binary matrix described above. Low values of entropy are assigned to words that are highly expected to appear in unique contexts, and high values of entropy are assigned to words whose context of appearance is highly unpredictable. The normalised entropy of each word i's probability vector was computed as follows:

$$H = \sum_{i=1}^{n} \frac{p(x_i)\log(p(x_i))}{\log(n)} \qquad (1)$$

where $p(x_i)$ denotes the proportion of co-occurrence of word i with word x, and n denotes the total number of words in the corpus. The term $\log(n)$ normalises the entropy, creating values between 0 and 1. An example of a low entropy word would be "tomato", which is more likely to co-occur with a restricted group of meal-related words. In contrast, an example of a high entropy word would be "boy", which is likely to co-occur with a more varied range of not-semantically related words.

5.2.3.2 Document Environment

In addition to computing contextual diversity and contextual distinctiveness based on word co-occurrences, we also calculated these two word measurements based on children's routines. To do this, we split our data files into five datasets sorted by topics: mealtime, playtime, nappy time, bath time, and bedtime. Contextual diversity for each word was then computed by counting the number of topics in which the word appears, with an appearance minimum of five times required to be considered as part of the context, since fixing this number to 1 makes most words fall into the maximum value of contextual diversity. Values of 10 and 20 gave similar results in our regression

models. Values of contextual diversity ranged between 0 and 5, with 0 representing low contextual diversity and 5 representing high contextual diversity. For entropy, we proceeded as above, but now $p(x_i)$ in equation (1) denotes the proportion of occurrences of word i with the topic x, and n denotes the total number of topics (=5).

5.2.4 Order of Word Acquisition

To calculate the order in which words are acquired by children, we counted the number of times each word was uttered by the children in our sample based on their CDI data. The assumption is that words that are acquired earlier are usually found across all vocabulary sizes, whereas words acquired later are found in large vocabularies only. Though the vocabulary data we collected is longitudinal, we only considered the vocabulary of the children at the end of the study, which varied sufficiently to capture the impact of the environment. Across children, the number of children who produced a given word was correlated with Kuperman et al.'s AoA (2012)($t(325)=-9.24$, $\varrho=-.46$), meaning that words produced by most children are also those that are produced earlier.

5.3 Results

We tested three contextual measures in predicting order of word acquisition (i.e., number of children that produced each word): frequency, contextual diversity and contextual distinctiveness. Since one of the aims of the present study is to determine the best context size that predicts word acquisition, we compared the predictive power of identical models whose word features were computed in different ways: from their lexical environment (i.e., lexical-diversity and lexical-distinctiveness) and from their document environment (i.e., document-diversity and document-distinctiveness). As explained in the method section, contextual distinctiveness was measured as the word's entropy, meaning that words with high entropy refer to words with low contextual distinctiveness, and vice versa. Before conducting a stepwise multiple regression, we first evaluated the possibility of a nonlinear pattern in the relation between each of the independent variables and the dependent variable. Polynomial regression analysis was conducted to predict order of word acquisition from each predictor in separate models, and for each model, an additional two versions were created where the predictor was either squared or cubed. We used the *lm* function and *anova* function in R (R Core Team, 2019) to conduct our regression analysis and compare the linear, quadratic and cubed versions for each predictor. Table 5.1 shows the BIC for these models. The best higher-

order models (as denoted in Table 5.1) were included in our hierarchical regression analysis. All our contextual measures are positively and highly correlated (see Table 5.2). We then investigated the relation between each word feature and word types (nouns, verbs, and adjectives). Table 5.3 shows the results obtained from non-parametric tests, where we compared each word type across word features. Nouns differ to verbs and adjectives across all word measures. Nouns are the word type that is produced the earliest (as denoted by the highest number of children that produced them), and with the lowest frequency, lexical-diversity and document-diversity, and with the highest lexical-distinctiveness and document-distinctiveness (i.e., lowest entropy). Adjectives and verbs have similar averages across variables, with the exception of frequency, where verbs presented higher frequency. Given the relation observed between word types and word features, we decided to control for word type in our regression models.

Table 5.1 Bayesian Information Criterion Scores for Fits to Contextual Measures

	Variable	Linear	Quadratic	Cubic
1	Frequency	1512	1493	**1490**
2	Lexical- diversity	1555	**1541**	1546
3	Document- diversity	1539	**1536**	1541
4	Lexical- distinctiveness	1559	**1560**	1561
5	Document - distinctiveness	1657	**1634**	1640

Note. The best-fitting model is marked in bold.

Table 5.2 Pearson Correlations among Contextual Measures

	Variable	1	2	3	*4*	*5*
1	Frequency	1.00				
2	Lexical- diversity	0.96	1.00			
3	Document- diversity	0.90	0.92	1.00		
4	Lexical- distinctiveness	0.87	0.89	0.81	1.00	
5	Document - distinctiveness	0.60	0.71	0.71	0.62	1.00

Table 5.4 shows the model results of our stepwise regression analysis. Our best multiple regression model included frequency, lexical-contextual diversity, and document-contextual distinctiveness (word type was included as a predictor in every step; frequency linear: $b = 1.56$, $SE = 0.18$, $p = < .001$; frequency quadratic: $b = -0.03$, $SE = 0.09$, $p = > .05$; frequency cubic: $b = -0.32$, $SE = 0.03$, $p = < .001$; diversity linear: $b = -0.57$, $SE = 0.22$, $p = < .01$; diversity quadratic: $b = -0.326$, $SE = 0.09$, $p = < .01$; distinctiveness linear: $b = -0.005$, $SE = 0.08$, $p = > .05$; distinctiveness quadratic: $b = -0.15$, $SE = 0.06$, $p = < .01$). A potential concern for comparing the model's coefficient estimates is predictor collinearity. Therefore, to interpret the effect of each of the best predictors (Step 5), we will refer to Figure 5.1. Figure 5.1 shows the 455 CDI words included in our analysis (each dot is a word) and the relation between the number of times a word is produced with our best predictors: word frequency, contextual diversity computed from lexical environment, and contextual distinctiveness computed from document environment (i.e., entropy). The graphs suggest that the more children produced a word, the higher its frequency and contextual diversity. In other words, the earliest words produced by children tend to be higher in frequency and contextual diversity, with the exception of a few cases. In contrast, contextual distinctiveness peaks near its middle values, meaning that children find it easier to produce words that are not too distinctive nor too diverse across contexts.

Table 5.3 Wilcoxon Rank Sum Test Comparing Nouns, Verb and Adjectives.

				Nouns vs verbs		Nouns vs adjectives		Verbs vs adjectives	
Variable	Mean Nouns	Mean Verbs	Mean Adj.	*W*	*d*	*W*	*d*	*W*	*d*
Word order	26.4	16.0	17.3	15332***	0.1	9863.5***	0.1	2582	-
Frequency	130.3	472.5	241.7	7936.5***	0.7	6283***	0.4	3450.5*	0.4
Lexical-diversity	355.9	446.1	436.9	6655***	0.9	4849***	0.6	3215.5	-
Document-diversity	2.9	4.3	4.3	7746***	0.7	5079.5***	0.6	2982.5	-
Lexical-distinctiveness	0.64	0.84	0.81	8098***	0.7	6228.5***	0.4	3363.5	-
Document-distinctiveness	0.70	0.87	0.91	7312.5***	0.8	3458.5***	0.9	2612	-

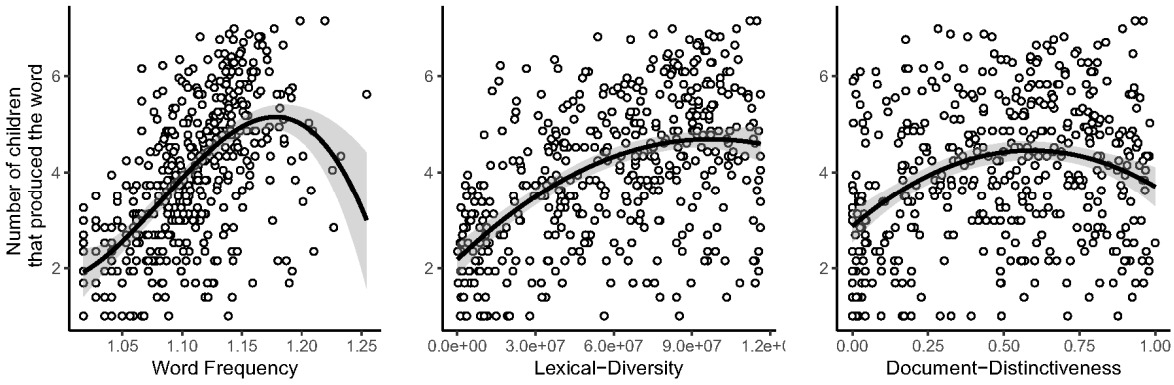
Note. All *p*-values were corrected using the BH method. Asterisks indicate a significant group difference: * $p < .05$; ** $p < .01$; *** $p < .001$.Values are not transformed. Word order is expressed as the number of children that produced each word per word class, where the higher the number the earlier acquisition of the word.

Table 5.4 Hierarchical Polynomial Regression Analysis Predicting Order of Word Acquisition

	Predictors	*F (df)*	*Adj.R^2*	*$R^2\Delta$*	*d*
Step 1	Frequency	96.24 (5,449)	0.51		0.51
Step 2	Frequency Lexical- diversity	74.73 (7,447)	0.53	0.02*** (vs. Step 1)	0.53
Step 3	Frequency Document-diversity	68.99 (7,447)	0.51	0.00 (vs. Step 1)	0.51
Step 4	Frequency Lexical- diversity Lexical- distinctiveness	58.66 (9,445)	0.53	0.00 (vs. Step 2)	0.53
Step 5	Frequency Lexical- diversity Document- distinctiveness	59.54 (9,445)	0.54	0.01* (vs. Step 2)	0.54

Note. All predictors were subject to transformations to comply with regression's assumptions. Tukey's Ladder of Powers approach was applied; in addition, word frequency was cubed, and lexical- diversity, lexical- distinctiveness, document- diversity and document- distinctiveness were squared. Word type was included as a predictor in each model. Asterisks indicate significance: * $p < .05$; ** $p < .01$; *** $p < .001$.

Figure 5.1 Scatterplot Depicting the Relation Between the Number of Children in the Sample that Produced each Word and the Words Features.



Note. Scatterplot depicting the relation between the number of children in the sample that produced each word and the words features, word frequency, lexical-diversity and document-distinctiveness (measure as entropy). Each dot represents a word. Predictors are transformed as described in Table 5.4. The number of times the word is produced represent the inverse order in which word are produced by children, and negatively correlated to AoA. Shadows around the curves represent 95% confidence intervals.

5.4 Discussion

Any measure of diversity naturally invites a measure of distinctiveness. In addition, these measures both invite questions about the scale at which measurements are made. The results of the present study demonstrate effects for both diversity and distinctiveness and further demonstrate that their difference may be at the level of scale. Both contextual diversity and contextual distinctiveness facilitate early word acquisition, but they appear to differ in the scale at which they act. Contextual diversity is more influential when computed at the lexical as opposed to document-level context. This may indicate that many of the results based on document-level measures would be enhanced by lexical level measures. The role of contextual distinctiveness is less clear but shows a slight bias for document (routine)-level as opposed to lexical-level distinctiveness. Our results suggest that words that are too high or too low in contextual distinctiveness may be more difficult to learn than words that experience more intermediate levels of distinctiveness.

The best contextual diversity predictor of word order was computed considering the immediate context of each unique word. This finding is suggestive of the important learning role of contextual diversity in the enrichment of the semantic representation of known words. In contrast, we found that the best measure of contextual distinctiveness to predict word acquisition is at larger contexts, up to the level of activities or routines. This finding may indicate that children need semi-consistent routines for word learning. This supports the idea suggested by Bruner (1985) that the consistency of transactional formats is essential for word learning in a way that children can predict their patterns of action. However, our finding adds to this statement that children also need small changes in their routines, possibly offering chances for word-referent mapping by mutual exclusivity. This hypothesis is consistent with previous findings from cross-situational learning that suggest that learning conditions in which the same words appear across contexts (i.e., Zipfian conditions, where most words are high-frequency words) allows the learner to discover the referent for the low-frequency words as they learn (Hendrickson & Perfors, 2019). Contextual distinctiveness computed considering the immediate context of each unique word appears to be also predictive of order of word acquisition by theoretically reinforcing the semantic representation of known words. This fits comfortably with empirical evidence that indicated that encountering novel words across redundant contexts helps to create stable semantic representations of the words (Johns, Dye, & Jones, 2016).

How does child-directed speech develop in relation to contextual diversity and contextual distinctiveness? Hills (2013) found that the speech of parents with younger children was less contextually-diverse than that of the parents with older children, even though children were still learning the most contextually diverse words earliest. These findings suggest that parents with children with small vocabularies make their speech more consistent and predictable by lowering the contextual diversity in their speech (i.e., possibly raising the lexical-level of contextual distinctiveness). This too may fit with Bruner's (1985) proposal that the consistency of interactional formats at early stages of language development guides the child in understanding the nature of language, but after this first stage, contextual diversity starts playing an important role in helping the child to map the words with their corresponding referents as well as enriching the words' concept through the exposure of their semantic features in different contexts.

We have computed our word features using aggregated data from various families. Similarly, we calculated order of word acquisition aggregating the vocabulary of

all the children in the sample. A better approach to answering our research questions would have been to compute these variables individually for each family and child. Unfortunately, that approach requires more data than we currently have leading us to aggregate environments and vocabularies across children. In addition, although we have covered the main common routines during infancy, there are potentially other routines and activities present in these early stages that we are not capturing, and that may also vary across families. As data collection continues to become easier, future work will be able to take these factors into account.

Chapter 6:
Discussion and Conclusions

6.1 Summary of the Findings

As well as mapping the heard word to the correct referent in the world, young learners gather semantic information about the word which promotes the formation of mental concepts. The present thesis aims to investigate the internal and external factors involved in the semantic enrichment of words' concepts and their influence on early lexical development. Specifically, this thesis focuses on *contextual diversity,* a word feature known to influence semantic development. To do this, this thesis examines 1) the association of the structural characteristics of the lexicon related to time of language onset (delay versus no-delay), 2) the cognitive processes involved in the semantic maturation of words prior to production, 3) the qualitative characteristics of the language environment that allows learning the meaning of words through contextual diversity, and 4) the different metrics of contextual diversity to determine its main role in early word learning. In what follows, I summarise the main findings reported in this thesis in relation to the investigated aspects that are enumerated above, and build connections between the findings.

Structure of the lexicon. In Chapter 2, children with and without delay produced similar proportions of words across most semantic categories, with some exceptions in which most differences reflected the age difference between the groups. Regarding syntactic classes, although children with language delay showed noun bias in their lexicons, a reduced noun-verb gap was noticed. Specifically, this gap seems to be driven by the higher proportion of verbs and the lower proportion of nouns produced by the delayed groups in comparison to the typical talker group. Even though one of the delayed groups had ASD, the differences observed might be related to common language learning mechanisms as they seem to affect both delayed groups similarly, although to different degrees. All children showed the tendency to acquire earlier those verbs whose meaning involved highly-social interactions. However, children with ASD showed this social-word bias to a significantly lower degree, indicating an association between ASD characteristics and word acquisition. Regarding the network structure of

lexicons examined in Chapter 3, late talkers showed lower connectedness in their noun vocabularies but higher connectedness in their verb vocabularies than typical talkers. Without any further investigation into the underlying cognitive processes that raised these puzzling differences, it would be sensible to ask whether there is an association between the observed reduced noun-verb production gap found in Chapter 2 and the disparity of the connectedness in noun and verb vocabularies found in Chapter 3. Finally, the results in Chapter 4 draw a relevant association between the semantic network-structure of the language environment and that of the children's vocabularies: the semantic associations that late talkers acquire from their language environments drive their vocabulary networks to have weaker connections at a global level.

Cognitive processes in semantic maturation. The best computational network growth model presented in Chapter 3 suggests that words that enter the receptive vocabulary experience a period of semantic maturation before moving into production. Nouns and verbs mature at different rates, with verbs requiring more semantic maturation than nouns. The results also indicated that typical talkers have more semantic maturation in the words they produce than late talkers, even though both talker groups showed similar attentional sensitivity towards contextual diversity. Further, the semantic maturation gap between nouns and verbs is significantly smaller for late talkers, meaning that verbs require less semantic maturation for late talkers than for typical talkers. This lower semantic verb maturation makes late talkers produce verbs earlier relative to typical talkers, which may explain the reduced noun-verb production gap in children with language delay found in Chapter 2.

Semantic richness in the language environment. In Chapter 4, the parental speech directed to children with language delay was found to be semantically poorer, with semantic richness operationalised as contextual diversity. In addition, the network structure of the speech directed to late talkers was found to be less well-connected. Crucially, the quality of the environment of the speech received by late bloomers mostly resembles that of typical talkers. Therefore, these findings provide a valid answer to the question posed in Chapter 3: a semantically poor language input might have generated the lower maturation in the words that late talkers produce.

The role of contextual diversity in early word learning. Contextual diversity measured as the number of other word types the word co-occurs with in the immediate context (i.e., within a small window size) is more predictive of order of word acquisition than

contextual diversity measured by co-occurrence in an ampler context (i.e., across routines or activities). This finding from Chapter 5 suggests that the main advantage of contextual diversity in word learning is to semantically enrich the concepts attached to words, as close semantic features and nuances of the word are more likely to be found near to the word in question. The fact that contextual diversity was computed utilising a window-context in all the studies presented in this thesis confirms the idea that semantic maturation is the core phenomena behind our findings. In addition, contextual distinctiveness (another contextual measure that refers to the consistency of the context in which the word is embedded) is also a valid predictor of word acquisition after frequency and contextual diversity; however, earlier words are associated with intermediate degrees of contextual distinctiveness. This finding suggests that, although consistency is good for lexical development, small variations in routines might facilitate the use of mutual exclusivity to figure out the referents for novel words.

6.2 Theoretical Implications

In what follows, I describe the five main theoretical implications of the research presented in this thesis for existing language acquisition accounts.

Word acquisition is not just mapping words with their referents. Most research that investigates the fundamental skills and processes needed for word learning regarding the behaviour of mapping a word with the correct referent as word learning. In other words, there is a general conception that there is a particular moment in life when a child makes the connection between what is heard and something in the world. This is the very moment when one would identify that the child has learned a new word. However, the child has experienced the referent many times before having made the word-referent mapping. Even after the child learns the referent of a word, the child remains silent, and it is not for some months that she produces it for the first time. This thesis has provided evidence that the preverbal semantic information that the child gathers by experiencing the word in different contexts seems to determine the type of words the child will produce. Therefore, the findings presented in this thesis implicate the redefinition of early word acquisition in which it is considered those experiences with the referent until the child maps the referent with the correct reference-word. This rationale makes the whole idea of word learning a complex one as anything can be encoded or mentally associated with a word: places, smells, other words, or even the language used at the moment of encoding. This thesis only covers the linguistic context,

but it has been shown that spatial and temporal contexts (i.e., where and when the child hears the word) are also encoded by the child along with the word, and these word-context consistencies act as learning facilitators (Roy et al., 2015). In sum, this thesis promotes the idea that word-referent mapping is just the tip of the iceberg of early word acquisition.

Semantic maturation promotes word production. This conclusion, derived from the findings presented in this thesis, proposes a further step towards a wider view of lexical acquisition. Accordingly, the first time the child produces a particular word is regarded as the result of encountering that word in a number of diverse situations. Previous studies that investigated the influence of learning novel words under different conditions of contextual diversity reported a facilitation effect in recognising and retrieving (Goldinger & Azuma, 2004; McDonald & Shillcock, 2001; Nelson & Shiffrin, 2006; Pexman et al., 2008; Adelman, Brown, & Quesada, 2006; Frances, Martin & Duñabeita, 2020). The cognitive advantages that entail a structural organisation based on semantic associations between words might well explain these facilitation effects. From a connectionist point of view, the activation of the representation of the word seems to be boosted by the size of the semantic cluster in which the word is embedded as well as the number of semantic links that the word has with other words in the lexicon. The explanation might be satisfactory in the case of adults and older children who have mastered the use of language. This is less clear when we talk about young children who have just started to produce their first words. The cognitive processes of comprehension and recognition are naturally related. When the child looks or points to the spoken word's referent, she is recognising the word. However, does this mean that the child 'understands' the word? Our best vocabulary growth model builds up a comprehension vocabulary based on the semantic enrichment of the mental word's concept that the child gains from encountering the words in different contexts. This means that there is not a binary answer to the question of whether the child understands the word or not. The phonological output representation of a concept, i.e. the 'spoken word', is just one of the many pieces that together constitute the concept in the child's mind. Theoretically, this would mean that for a child to be able to spontaneously retrieve this piece of information about the concept (its spoken form), the more semantic information attached to the concept, the easier it is to be retrieved. Therefore, following this assumption, building the mental word's concept seems critical for typical word production.

Semantically-poor environments might impede normal lexical development. This thesis suggests that semantically-poor language environments were associated with children who showed early language delay. Previous studies have already suggested the link between the diversity of the word types in parental speech and early language acquisition (Rowe, 2013; Thomas & Knowland, 2014; Newman, Rowe, & Ratner; 2016). Many studies have focused on how children exploit the distributional characteristics in child-directed speech for word learning (e.g., Mintz, Newport, & Bever, 2002; Weisleder & Waxman, 2010; Mccauley, Monaghan & Christiansen, 2015). The findings derived from Chapter 4 suggest that the semantic quality of the linguistic input might influence word learning, not only in the type of to-be-learned words but also when words will be acquired. Furthermore, the semantic links that late talkers learn from their respective language environments are sufficient to guarantee that their productive vocabularies will show a less advantageous network structure that could be related to a weaker representation of the concepts/words. This is important because they way in which words are organised in the mental lexicons have an impact on language processing (e.g., Goldinger & Azuma, 2004; Adelman, Brown, & Quesada, 2006). The findings open a new avenue of investigation in word acquisition research that could explore the effects of the amount of semantic information conveyed by the distributional properties of the language environment on how children process words.

Investigating external and internal factors to get the whole picture. The majority of studies that aimed to elucidate the origin of early language delay have primarily focused on children's word learning abilities. Although these internal aspects are important to examine, the reality is that we do not know much about the home learning environment that the participant children came from, and what they had experienced previously. As per this thesis, with the sole information provided from the models and the characteristics of the children's vocabularies, one could have concluded that language delay might be associated with children's inability to use the information embedded in the semantic structure of child-directed speech. However, the examination of the semantic quality of the speech of late talkers' caregivers changed the perspective of the issue completely. Therefore, the accuracy in answering a research question can be greatly enhanced by the use of a holistic and naturalistic approach to investigating the phenomena. This implicates that some learning mechanisms that are thought to be linked to late language onset would be worth revisiting to further examine the characteristics of the children's natural learning environment.

Dimensional versus categorical view of language delay. The examination of the characteristics of the vocabulary of children with language delay can provide clues on the learning mechanisms that might have operated atypically to reach such lexical profiles. However, differences in the lexical profile might not be enough evidence to discern between delay and deviance. The dimensional account defends the idea that all children possess the same learning tools that makes language acquisition possible. At the same time, children's skills might vary within a continuum, which implies that lower skills in using some learning tools might result in a delay. In contrast, the categorical account assumes that language delay is the result of some atypical learning mechanisms that are not normally present in children with typical language development. For example, it is expected that young children extend known names for referents to novel instances with the same shape, but late talkers extend known words to novel instances with the same texture instead (Jones, 2003). Similarly, it was hypothesised that late talkers might be attracted to produce 'oddball' words as opposed to well-connected words in the learning environment (Beckage et al., 2011). These are some examples of how one could label word learning as 'atypical' or 'deviant'. Many differences in the proportion of semantic categories and syntactic categories described in Chapter 2 suggested an age effect. This is, the direction in the proportion differences reflected the extent of the language delay between the groups. Also, the data in Chapter 2 do not relate to processes whereby the children acquired the words (i.e., any learning mechanism like shape bias), which could potentially confirm a deviant language development. The study in Chapter 2 also surveyed a new potential word bias in children: the social-word bias. Although children with ASD acquired fewer high-social words than neurotypical children, they also showed the same growth curve or social-word bias. Likewise, the differences identified in Chapter 3 in the semantic structure of children with and without language delay might be well explained by a spectrum of abilities involved in the normal cognitive processes of semantic maturation, as the best model disclosed. Overall, the results presented in this thesis support the dimensional account of language delay since differences in the proficiency of typical learning mechanisms can explain the differences between the talker groups.

6.3 Practical Implications

One of the conclusions of this thesis is that the quality of the semantic richness in child-directed speech might have an effect on lexical development. The logic is that

poor parental input might limit the mental formation of rich concepts of words, which is suggested to determine when a word will be produced. In part, this evidence means good news for practitioners, who can utilise this information to create potential effective interventions. The intervention could take the form of the practitioner working directly with the families to teach them how to enrich their daily interactions with their child. Two main aspects should be taken into account in the design of the intervention. The first aspect is the selection of target words. These words should be high in contextual diversity because their numerous semantic links with other words make them ideal to be introduced in many distinct contexts, where its associates await. Second, since semi-consistent contexts are also necessary for word acquisition, not many target words should be introduced at the same time in the same context. Semi-consistent contexts could also encourage word-referent mapping by mutual exclusivity in the case that the child does not know the referent for the novel word yet. An intervention based on contextual diversity has some advantages: 1) it is a non-invasive intervention as the therapist does not work directly with the child, 2) since the child spends most of her infancy surrounded by her caregivers, changes in the daily language input might lead to a greater effect, 3) the treatment can start at any time, being a natural approach to carry out during the 'wait and see' period. However, on the downside, meta-analysis studies reported that when the intervention is provided by the caregivers the effect of the treatment is lower compared to when the experimenter gave the treatment (Marulis and Neuman, 2010). Therefore, parents might need significant training to maximise the benefits of a contextual diversity intervention.

6.4 Future Directions

Despite the efforts made to elucidate the impact of contextual diversity in word acquisition in this thesis, more work is needed to advance the understanding of the role of this word feature. The following are some general suggestions for future research in this field.

There are many experimental studies that have exposed the benefits of encountering a word in different contexts. Adults and older children have been the target population to test this hypothesis. It is noted that empirical studies in contextual diversity with infants are lacking. The approach of exploring the natural environment of children has many ecological advantages; however, it is best to complement these

observations with the development of experiments in controlled environments to confirm the causality suggested by the results produced in this thesis.

Another valuable contribution to the literature on semantic maturation would be the examination of the impact of the different types of semantic features on word acquisition. In this thesis, we considered the semantic links the word has with other words in the lexicon, mostly with other nouns and verbs. This contextual diversity measure assumes that all semantic links between words have equal value, however this might not be the case. For instance, the link between 'giraffe' and 'neck' might be stronger than 'giraffe' and 'big'. Therefore, the use of recalculated values of contextual diversity by 'weighting' the links between words might show some new interesting patterns in the network trajectories.

After the investigation of the children's actual language environment, I concluded that the lower semantic maturation of the words learned by children with delayed onset might be associated with a poor linguistic input. Intervention studies are essential to confirm that late talkers are able to learn typically if the language environment is optimal for accessing the semantic features of words. Importantly, these intervention studies should consider the age and the vocabulary size of the participant infants given the findings derived from Chapter 4.

Finally, the models presented in Chapter 3 utilised the same learning environment as an input. Given that the dissimilarity of the semantic richness in child-directed speech between the families of late talkers and typical talkers, future research should consider utilising different linguistic inputs for vocabulary growth models, something I am planning to carry out in the near future for the models in Chapter 3.

6.5 Closing Remarks

The present thesis is the result of an effort made to provide a comprehensive understanding of the internal and external factors involved in the semantic enrichment of the word's concept and its relation with lexical development. I see the investigation of the formation of the preverbal comprehension lexicon as a broad field that has just started to be explored. I hope that this thesis will encourage and inspire future research exploring semantic maturation in early word acquisition. Particularly, I would be grateful if the findings from this thesis or from the studies that will follow could advise in the

design of effective interventions that help young children to promptly recover or avoid early language delay.

References

Adelman, J. S., Brown, G. D., & Quesada, J. F. (2006). Contextual diversity, not word frequency, determines word-naming and lexical decision times. *Psychological Science*, *17*(9), 814-823. https://doi.org/10.1111/j.1467-9280.2006.01787.x

Akhtar, N., & Montague, L. (1999). Early lexical acquisition: The role of cross-situational learning. *First Language*, *19*(57), 347-358. https://doi.org/10.1177/014272379901905703

Alcock, K., Meints, K., & Rowland, C. (2020). *The UK communicative development inventories: Words and gestures*. J&R Press.

Alt, M., Mettler, H. M., Erikson, J. A., Figueroa, C. R., Etters-Thomas, S. E., Arizmendi, G. D., & Oglivie, T. (2020). Exploring input parameters in an expressive vocabulary treatment with late talkers. *Journal of Speech, Language, and Hearing Research*, *63*(1), 216-233. https://doi.org/10.1044/2019_jslhr-19-00219

Alt, M., Meyers, C., Oglivie, T., Nicholas, K., & Arizmendi, G. (2014). Cross-situational statistically based word learning intervention for late-talking toddlers. *Journal of Communication Disorders*, *52*, 207-220. https://doi.org/10.1016/j.jcomdis.2014.07.002

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington DC: American Psychiatric Association.

Amsel, B. D. (2011). Tracking real-time neural activation of conceptual knowledge using single-trial event-related potentials. *Neuropsychologia*, *49*(5), 970-983. https://doi.org/10.1016/j.neuropsychologia.2011.01.003

Anderson, J. R., & Milson, R. (1989). Human memory: An adaptive perspective. *Psychological Review*, *96*(4), 703. https://doi.org/10.1037/0033-295x.96.4.703

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science*, *2*(6), 396-408. https://doi.org/10.1111/j.1467-9280.1991.tb00174.x

Antonucci, S. M., & Alt, M. (2011). A lifespan perspective on semantic processing of concrete concepts: does a sensory/motor model have the potential to bridge the gap? *Cognitive, Affective, & Behavioral Neuroscience*, *11*(4), 551-572. https://doi.org/10.3758/s13415-011-0053-y

Barber, H. A., Otten, L. J., Kousta, S. T., & Vigliocco, G. (2013). Concreteness in word processing: ERP and behavioral effects in a lexical decision task. *Brain and Language*, *125*(1), 47-53. https://doi.org/10.1016/j.bandl.2013.01.005

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. *The Crosslinguistic Study of Sentence Processing*, *3*, 73-112.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., … Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, 21(1), 85–123. https://doi.org/10.1017/s0305000900008680

Beckage, N., Smith, L., & Hills, T. (2011). Small Worlds and Semantic Network Growth in Typical and Late Talkers. *PLoS ONE*, *6*(5). https://doi.org/10.1371/journal.pone.0019348

Benedict, H. (1979). Early lexical development: Comprehension and production. *Journal of Child Language*, *6*(2), 183-200. https://doi.org/10.1017/s0305000900002245

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

Bilson, S., Yoshida, H., Tran, C. D., Woods, E. A., & Hills, T. T. (2015). Semantic Facilitation in Bilingual First Language Acquisition. *Cognition*, *140*, 122–134. http://doi.org/10.1016/j.cognition.2015.03.013

Bishop, D. V. (2006). What causes specific language impairment in children? *Current directions in psychological science*, *15*(5), 217-221. https://doi.org/10.1111/j.1467-8721.2006.00439.x

Bishop, D. V., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry*, *31*(7), 1027-1050. https://doi.org/10.1111/j.1469-7610.1990.tb00844.x

Bishop, D. V., & Edmundson, A. (1987). Language-impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders*, *52*(2), 156-173. https://doi.org/10.1044/jshd.5202.156

Bishop, D. V., Holt, G., Line, E., McDonald, D., McDonald, S., & Watt, H. (2012). Parental phonological memory contributes to prediction of outcome of late talkers from 20 months to 4 years: a longitudinal study of precursors of specific language impairment. *Journal of Neurodevelopmental Disorders*, *4*(1), 1-12. https://doi.org/10.1186/1866-1955-4-3

Bishop, D. V., North, T., & Donlan, C. (1995). Genetic basis of specific language impairment: Evidence from a twin study. *Developmental Medicine & Child Neurology*, *37*(1), 56-71. https://doi.org/10.1111/j.1469-8749.1995.tb11932.x

Bishop, S. L., Farmer, C., & Thurm, A. (2015). Measurement of nonverbal IQ in autism spectrum disorder: scores in young adulthood compared to early childhood. *Journal of Autism and Developmental Disorders*, *45*(4), 966-974. https://doi.org/10.1007/s10803-014-2250-3

Bishop, S. L., Guthrie, W., Coffing, M., & Lord, C. (2011). Convergent validity of the Mullen Scales of Early Learning and the differential ability scales in children with autism spectrum disorders. *American Journal on Intellectual and Developmental Disabilities*, *116*(5), 331-343. https://doi.org/10.1352/1944-7558-116.5.331

Blythe, H. I., Liversedge, S. P., Joseph, H. S., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision Research*, *49*(12), 1583-1591. https://doi.org/10.1016/j.visres.2009.03.015

Braginsky, M., Yurovsky, D., Marchman, V. A., & Frank, M. C. (2019). Consistency and variability in children's word learning across languages. *Open Mind*, 1-16. https://doi.org/10.31234/osf.io/cg6ah

Brown, P. (2008). Verb specificity and argument realization in Tzeltal child language. *Crosslinguistic Perspectives on Argument Structure: Implications for Learnability*, 167-189.

Bruckner, C., Yoder, P., Stone, W., & Saylor, M. (2007). Receptive Vocabulary Scale Can Be Can Be Improved: Differential Item Functioning Between Toddlers with Autism Spectrum Disorders and Typically Developing Infants. *Journal of Speech, Language, and Hearing Research*, 50(12), 1631–1638.

Bruner, J. (1985). Child's talk: Learning to use language. *Child Language Teaching and Therapy*, *1*(1), 111-114. https://doi.org/10.1177/026565908500100113

Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46(3), 904-911. https://doi.org/10.3758/s13428-013-0403-5

Cancho, R. F. I., & Solé, R. V. (2001). The small world of human language. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1482), 2261-2265. https://doi.org/10.1098/rspb.2001.1800

Capone, N. C., & McGregor, K. K. (2005). The effect of semantic representation on toddlers' word retrieval. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2005/102)

Capute, A. J., & Accardo, P. J. (1996). A neurodevelopmental perspective on the continuum of developmental disabilities. *Developmental Disabilities in Infancy and Childhood*, *1*, 1-22.

Caselli, M. C., & Casadio, P. (1995). *Il primo vocabolario del bambino: guida all'uso del questionario MacArthur per la valutazione della comunicazione e del linguaggio nei primi anni di vita* (Vol. 5). FrancoAngeli.

Ceponiene, R., Lepistö, T., Shestakova, A., Vanhala, R., Alku, P., Näätänen, R., & Yaguchi, K. (2003). Speech-sound-selective auditory impairment in children with autism: they can perceive but do not attend. *Proceedings of the National Academy of Sciences of the United States of America*, 100(9), 5567–5572. https://doi.org/10.1073/pnas.0835631100

Chan, K. Y., & Vitevitch, M. S. (2009). The influence of the phonological neighborhood clustering coefficient on spoken word recognition. *Journal of Experimental Psychology: Human Perception and Performance,* 35(6), 1934. https://doi.org/10.1037/a0016902

Charman, T., Drew, A., Baird, C., & Baird, G. (2003). Measuring early language development in preschool children with autism spectrum disorder using the MacArthur Communicative Development Inventory (Infant Form). *Journal of Child Language*, 30(1), 213-236. https://doi.org/10.1017/s0305000902005482

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The Social Motivation Theory of Autism Introduction: Social motivation and social cognition, two competing. *Trends in Cognitive Sciences*, 16(4), 231–239. https://doi.org/10.1016/j.tics.2012.02.007

Clark, M. L., Barbaro, J., & Dissanayake, C. (2017). Continuity and change in cognition and autism severity from toddlerhood to school age. *Journal of Autism and Developmental Disorders*, *47*(2), 328-339. https://doi.org/10.1007/s10803-016-2954-7

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Collins, A. M., & Quillian, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, *8*(2), 240-247. https://doi.org/10.1016/s0022-5371(69)80069-1

Colunga, E., & Sims, C. E. (2017). Not only size matters: Early-talker and late-talker vocabularies support different word-learning biases in babies and networks. *Cognitive Science*, 41, 73–95. https://doi.org/10.1111/cogs.12409

Conner, J., Kelly-Vance, L., Ryalls, B., & Friehe, M. (2014). A play and language intervention for two-year-old children: Implications for improving play skills and language. *Journal of Research in Childhood Education*, *28*(2), 221-237. https://doi.org/10.1080/02568543.2014.883452

Conti-Ramsden, G., Botting, N., & Faragher, B. (2001). Psycholinguistic markers for specific language impairment (SLI). *Journal of Child Psychology and Psychiatry*, *42*(6), 741-748. https://doi.org/10.1111/1469-7610.00770

Csárdi, G., & Nepusz, T. (2006). The Igraph Software Package for Complex Network Research. *InterJournal Complex Systems*, *1695*, 1695.

Csillery, K., Francois, O. and Blum, M. G. B. (2012), abc: an R package for approximate Bayesian computation (ABC). *Methods in Ecology and Evolution* 3: 475-479. https://doi.org/1010.1111/j.2041-210X.2011.00179.x

Dale, P. S., & Goodman, J. C. (2005). Commonality and individual differences in vocabulary growth. *Beyond nature-nurture: Essays in honor of Elizabeth Bates*, 41-78.

De Houwer, A., Bornstein, M. H., & Leach, D. B. (2005). Assessing early communicative ability: A cross-reporter cumulative score for the MacArthur CDI. *Journal of Child Language*, *32*(4), 735. https://doi.org/10.1017/s0305000905007026

Dennis, S., & Humphreys, M. S. (2001). A context noise model of episodic word recognition. *Psychological Review*, *108*(2), 452. https://doi.org/10.1037/0033-295x.108.2.452

Desmarais, C., Sylvestre, A., Meyer, F., Bairati, I., & Rouleau, N. (2008). Systematic review of the literature on characteristics of late-talking toddlers. *International Journal of Language & Communication Disorders*, *43*(4), 361-389. https://doi.org/10.1080/13682820701546854

Dhillon, R. (2010). Examining the 'Noun Bias': A structural approach. *University of Pennsylvania Working Papers in Linguistics*, *16*(1), 7.

D'Odorico, L., & Jacob, V. (2006). Prosodic and lexical aspects of maternal linguistic input to late-talking toddlers. *International Journal of Language & Communication Disorders*, *41*(3), 293-311. https://doi.org/10.1080/13682820500342976

D'odorico, L., Assanelli, A., Franco, F., & Jacob, V. (2007). A follow-up study on Italian late talkers: Development of language, short-term memory, phonological awareness, impulsiveness, and attention. *Applied Psycholinguistics*, *28*(1), 157-169.

Dollaghan, C. A. (2004). Taxometric analyses of specific language impairment in 3-and 4-year-old children. *Journal of Speech, Language & Hearing Research*, 47(4), 464-475. https://doi.org/10.1017/s0142716407070087

Dominey, P. F., & Dodane, C. (2004). Indeterminacy in language acquisition: the role of child directed speech and joint attention. *Journal of Neurolinguistics*, *17*(2-3), 121-145. https://doi.org/10.1016/s0911-6044(03)00056-3

Dykstra, J. R., Boyd, B. A., Watson, L. R., Crais, E. R., & Baranek, G. T. (2012). The impact of the Advancing Social-communication And Play (ASAP) intervention on preschoolers with autism spectrum disorder. *Autism*, 16(1), 27-44. https://doi.org/10.1177/1362361311408933

Eigsti, I. M., De Marchena, A. B., Schuh, J. M., & Kelley, E. (2011). Language acquisition in autism spectrum disorders: A developmental review. *Research in Autism Spectrum Disorders*, 5(2), 681–691. https://doi.org/10.1016/j.rasd.2010.09.001

Eimas, P. D., & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development*, *65*(3), 903-917. https://doi.org/10.2307/1131427

Ellis Weismer, S. (2017). Typical talkers, late talkers, and children with specific language impairment: A language endowment spectrum? In *Language disorders from a developmental perspective* (pp. 83-101). Psychology Press. https://doi.org/10.4324/9781315092041-3

Ellis Weismer, S., Gernsbacher, M. A., Stronach, S., Karasinski, C., Eernisse, E. R., Venker, C. E., & Sindberg, H. (2011). Lexical and grammatical skills in toddlers on the autism spectrum compared to late-talking toddlers. *Journal of Autism and Developmental Disorders*, 41(8), 1065–1075. doi.org/10.1007/s10803-010-1134-4

Ellis, E. M., Borovsky, A., Elman, J. L., & Evans, J. L. (2015). Novel word learning: An eye-tracking study. Are 18-month-old late talkers really different from their typical peers? *Journal of Communication Disorders*, *58*, 143-157. https://doi.org/10.1016/j.jcomdis.2015.06.011

Engelthaler, T., & Hills, T. T. (2017). Feature biases in early word learning: network distinctiveness predicts age of acquisition. *Cognitive Science*, *41*, 120-140. https://doi.org/10.1111/cogs.12350

Eriksson, M., Marschik, P. B., Tulviste, T., Almgren, M., Pérez Pereira, M., Wehberg, S., ... & Gallego, C. (2012). Differences between girls and boys in emerging language skills: Evidence from 10 language communities. *British journal of developmental psychology*, *30*(2), 326-343. https://doi.org/10.1111/j.2044-835X.2011.02042.x

Evert, S. (2008). Corpora and Collocations. *Corpus Linguistics. An International Handbook*, 1–53. http://doi.org/10.1515/9783110213881.2.1212

Fagundes, N. J., Ray, N., Beaumont, M., Neuenschwander, S., Salzano, F. M., Bonatto, S. L., & Excoffier, L. (2007). Statistical evaluation of alternative models of human evolution. *Proceedings of the National Academy of Sciences*, 104(45), 17614-17619. https://doi.org/10.1073/pnas.0708280104

Feijoo, S., Muñoz, C., Amadó, A., & Serrat, E. (2017). When meaning is not enough: Distributional and semantic cues to word categorization in child directed speech. *Frontiers in Psychology*, *8*, 1242. https://doi.org/10.3389/fpsyg.2017.01242

Feldman, H. M., Dale, P. S., Campbell, T. F., Colborn, D. K., Kurs-Lasky, M., Rockette, H. E., & 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. https://doi.org/10.1111/j.1467-8624.2005.00882.x

Feldman, H. M., Dollaghan, C. A., Campbell, T. F., Kurs-Lasky, M., Janosky, J. E., & Paradise, J. L. (2000). Measurement properties of the MacArthur Communicative Development Inventories at ages one and two years. *Child Development*, *71*(2), 310-322. https://doi.org/10.1111/1467-8624.00146

Fenson, L., Marchman, V.A., Thal, D.J., Dale, P.S., Reznick, J. S, & Bates, E. (2007). *MacArthur-Bates communicative development inventories*. Baltimore, MD: Paul H. Brookes Publishing Company.

Fenson, L., Bates, E., Dale, P., Goodman, J., Reznick, J. S., & Thal, D. (2000). Reply: Measuring variability in early child language: Don't shoot the messenger. *Child Development*, *71*(2), 323-328.

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*, i-185. https://doi.org/10.1111/1467-8624.00147

Fenson, L., Dale, P. S., Reznick, J. S., Thal, D., Bates, E., Hartung, J. P., Pethick, S. & Reilly, J. S. (1993). *MacArthur Communicative Development Inventory: users guide and technical manual.* San Diego, CA: Singular Publishing Company.

Fernald, A., & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically developing and late-talking toddlers. *Child Development*, *83*(1), 203-222. https://doi.org/10.1111/j.1467-8624.2011.01692.x

Field, C., Allen, M. L., & Lewis, C. (2016). Are children with autism spectrum disorder initially attuned to object function rather than shape for word learning? *Journal of Autism and Developmental Disorders*, 46, 1210–1219. https://doi.org/10.1007/s10803-015-2657-5

Fisher, E. L. (2017). A systematic review and meta-analysis of predictors of expressive-language outcomes among late talkers. *Journal of Speech, Language, and Hearing Research*, *60*(10), 2935-2948. https://doi.org/10.1044/2017_jslhr-l-16-0310

Forster, K. I., & Chambers, S. M. (1973). Lexical access and naming time. *Journal of Verbal Learning and Verbal Behavior*, *12*(6), 627-635. https://doi.org/10.1016/s0022-5371(73)80042-8

Fourtassi, A., Bian, Y., & Frank, M. C. (2019). The growth of children's semantic and phonological networks: insight from 10 languages. *Cognitive Science*, *44*, e12847. https://doi.org/10.31234/osf.io/37npj

Fraïsse, C., Roux, C., Gagnaire, P. A., Romiguier, J., Faivre, N., Welch, J. J., & Bierne, N. (2018). The divergence history of European blue mussel species reconstructed

from Approximate Bayesian Computation: The effects of sequencing techniques and sampling strategies. *PeerJ*, 6, e5198. https://doi.org/10.7717/peerj.5198

Frances, C., Martin, C. D., & Duñabeitia, J. A. (2020). The effects of contextual diversity on incidental vocabulary learning in the native and a foreign language. *Scientific reports*, *10*(1), 1-11. https://doi.org/10.1038/s41598-020-70922-1

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. https://doi.org/10.1017/s0305000916000209

Fritz, C. O., Morris, P. E., & Richler, J. J. (2012). Effect size estimates: current use, calculations, and interpretation. *Journal of Experimental Psychology: General*, *141*(1), 2. https://doi.org/10.1037/a0024338

Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child Development*, 988-998. https://doi.org/10.2307/1128738

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *Center for the Study of Reading Technical Report; no. 257*.

Gernsbacher, M. A., Geye, H. M., & Ellis Weismer, S. (2005). The role of language and communication impairments within autism. *Language Disorders and Developmental Theory*, 73-93. https://doi.org/10.1075/tilar.4.06ger

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*(2), 135-176. https://doi.org/10.1016/s0010-0277(99)00036-0

Girolametto, L., Bonifacio, S., Visini, C., Weitzman, E., Zocconi, E., & Pearce, P. S. (2002). Mother-child interactions in Canada and Italy: Linguistic responsiveness to late-talking toddlers. *International Journal of Language & Communication Disorders*, *37*(2), 153-171. https://doi.org/10.1080/13682820110116794

Girolametto, L., Pearce, P. S., & Weitzman, E. (1996). Interactive focused stimulation for toddlers with expressive vocabulary delays. *Journal of Speech, Language, and Hearing Research*, *39*(6), 1274-1283. https://doi.org/10.1044/jshr.3906.1274

Girolametto, L., Weitzman, E., Wiigs, M., & Pearce, P. S. (1999). The relationship between maternal language measures and language development in toddlers with expressive vocabulary delays. *American Journal of Speech-Language Pathology*, *8*(4), 364-374. https://doi.org/10.1044/1058-0360.0804.364

Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language*, *17*(1), 171-183. https://doi.org/10.1017/s0305000900013167

Goldinger, S. D., & Azuma, T. (2004). Episodic memory reflected in printed word naming. *Psychonomic Bulletin & Review*, 11(4), 716-722. https://doi.org/10.3758/bf03196625

Goldin-Meadow, S., Seligman, M. E., & Gelman, R. (1976). Language in the two-year old. *Cognition*, 4(2), 189-202. https://doi.org/10.1016/0010-0277(76)90004-4

Goodman, J. C., Dale, P. S., & Li, P. (2008). Does frequency count? Parental input and the acquisition of vocabulary. *Journal of Child Language*, *35*(3), 515-531. https://doi.org/10.1017/s0305000907008641

Gopnik, A., & Sobel, D. M. (2000). Detecting blickets: How young children use information about novel causal powers in categorization and induction. *Child Development*, *71*(5), 1205-1222. https://doi.org/10.1111/1467-8624.00224

Grossheinrich, N., Kademann, S., Bruder, J., Bartling, J., & Von Suchodoletz, W. (2010). Auditory sensory memory and language abilities in former late talkers: a mismatch negativity study. *Psychophysiology*, *47*(5), 822-830. https://doi.org/10.1111/j.1469-8986.2010.00996.x

Hadley, P. A., Rispoli, M., & Hsu, N. (2016). Toddlers' verb lexicon diversity and grammatical outcomes. *Language, Speech, and Hearing Services in Schools*, *47*(1), 44-58. https://doi.org/10.1044/2015_lshss-15-0018

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language*, *27*(3), 689-705. https://doi.org/10.1017/s0305000900004414

Hani, H. B., Gonzalez-Barrero, A. M., & Nadig, A. S. (2013). Children's referential understanding of novel words and parent labeling behaviors: Similarities across children with and without autism spectrum disorders. *Journal of Child Language*, *40*(5), 971-1002. https://doi.org/10.1017/s0305000912000426

Happé, F. G. E., & Booth, R. D. L. (2008). The power of the positive: Revisiting weak coherence in autism spectrum disorders. *Quarterly Journal of Experimental Psychology*, 61, 50-63. https://doi.org/10.1080/17470210701508731

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.

Hartig, F., Calabrese, J. M., Reineking, B., Wiegand, T., & Huth, A. (2011). Statistical inference for stochastic simulation models–theory and application. *Ecology Letters*, *14*(8), 816-827. https://doi.org/10.1111/j.1461-0248.2011.01640.x

Hawa, V. V., & Spanoudis, G. (2014). Toddlers with delayed expressive language: An overview of the characteristics, risk factors and language outcomes. *Research in Developmental Disabilities*, *35*(2), 400-407. https://doi.org/10.1016/j.ridd.2013.10.027

Hendrickson, A. T., & Perfors, A. (2019). Cross-situational learning in a Zipfian environment. *Cognition*, 189, 11-22. https://doi.org/10.1016/j.cognition.2019.03.005

Hills, T. (2013). The company that words keep: comparing the statistical structure of child- versus adult-directed language. *Journal of Child Language*, 40(3). 586-604. http://doi.org/10.1017/S0305000912000165

Hills, T. T., Maouene, J., Riordan, B., & Smith, L. B. (2010). The Associative Structure of Language: Contextual Diversity in Early Word Learning. *Journal of Memory and Language*, *63*(3), 259–273. http://doi.org/10.1016/j.jml.2010.06.002

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009a, May). Longitudinal Analysis of Early Semantic Networks Preferential Attachment or Preferential Acquisition? *Psychological Science,* 20, 729-739. https://doi.org/10.1111/j.1467-9280.2009.02365.x

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. (2009b, September). Categorical Structure Among Shared Features in Networks of Early-Learned Nouns. *Cognition*, *112*(3), 381–396. https://doi.org/10.1016/j.cognition.2009.06.002

Hilton, M., Twomey, K. E., & Westermann, G. (2019). Taking their eye off the ball: How shyness affects children's attention during word learning. *Journal of Experimental Child Psychology*, *183*, 134-145. https://doi.org/10.1016/j.jecp.2019.01.023

Hirsh-Pasek, K., & Golinkoff, R. M. (Eds.). (2010). *Action meets word: How children learn verbs.* Oxford University Press.

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development*, *73*(2), 418-433. https://doi.org/10.1111/1467-8624.00415

Hoffman, P., Ralph, M. A. L., & Rogers, T. T. (2013). Semantic diversity: A measure of semantic ambiguity based on variability in the contextual usage of words. *Behavior Research Methods*, 45(3), 718-730. https://doi.org/10.3758/s13428-012-0278-x

Holcomb, P. J., Grainger, J., & O'rourke, T. (2002). An electrophysiological study of the effects of orthographic neighborhood size on printed word perception. *Journal of Cognitive Neuroscience*, *14*(6), 938-950. https://doi.org/10.1162/089892902760191153

Horvath, S., McDermott, E., Reilly, K., & Arunachalam, S. (2018). Acquisition of verb meaning from syntactic distribution in preschoolers with autism spectrum disorder. *Language, Speech, and Hearing Services in Schools*, *49*(3S), 668-680. https://doi.org/10.1044/2018_lshss-stlt1-17-0126

Horvath, S., Rescorla, L., & Arunachalam, S. (2018). Semantic features of early verb vocabularies. In K. Syrett & S. Arunachalam (Eds.), *Semantics in language acquisition* (Vol. 24, pp. 67–94). Philadelphia, PA: John Benjamins Publishing Company.

Horvath, S., Rescorla, L., & Arunachalam, S. (2019). The syntactic and semantic features of two-year-olds' verb vocabularies: A comparison of typically developing

children and late talkers. *Journal of Child Language*, *46*(3), 409-432. https://doi.org/10.1017/s0305000918000508

Hsiao, Y., & Nation, K. (2018). Semantic diversity, frequency and the development of lexical quality in children's word reading. *Journal of Memory and Language*, *103*, 114-126. https://doi.org/10.31234/osf.io/p9xk4

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. https://doi.org/10.1037/0012-1649.27.2.236

Huttenlocher, J., Smiley, P., & Charney, R. (1983). Emergence of action categories in the child: Evidence from verb meanings. *Psychological Review*, *90*(1), 72. https://doi.org/10.1037/0033-295x.90.1.72

Ingersoll, B., & Schreibman, L. (2006). Teaching reciprocal imitation skills to young children with autism using a naturalistic behavioral approach: Effects on language, pretend play, and joint attention. *Journal of Autism and Developmental Disorders*, *36*(4), 487. https://doi.org/10.1007/s10803-006-0089-y

Irwin, J. R., Carter, A. S., & Briggs-Gowan, M. J. (2002). The social-emotional development of "late-talking" toddlers. *Journal of the American Academy of Child & Adolescent Psychiatry*, 41(11), 1324-1332. https://doi.org/10.1097/00004583-200211000-00014

Jiménez, E., Haebig, E., & Hills, T. T. (2020). Identifying areas of overlap and distinction in early lexical profiles of children with autism spectrum disorder, late talkers, and typical talkers. *Journal of Autism and Developmental Disorders*, 1-17. https://doi.org/10.1007/s10803-020-04772-1

Jiménez, E., & Hills, T. T. (2017). Network Analysis of a Large Sample of Typical and Late Talkers. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society CogSci*.

Jiménez, E., & Hills, T (n.d.). Semantic Maturation During the Comprehension-Expression Gap in Late and Typical Talkers. *Psychological Science*.

Johns, B. T., Dye, M., & Jones, M. N. (2016). The Influence of Contextual Variability in Word Learning. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-015-0980-7

Jones, M. N., Johns, B. T., & Recchia, G. (2012). The role of semantic diversity in lexical organization. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *66*(2), 115. https://doi.org/10.1037/a0026727

Jones, M. N., & Mewhort, D. J. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review*, 114(1), 1. https://doi.org/10.1037/0033-295x.114.1.1

Jones, S. S. (2003). Late talkers show no shape bias in a novel name extension task. *Developmental Science*, *6*(5), 477-483. https://doi.org/10.1111/1467-7687.00304

Joseph, H., Nation, K., & Liversedge, S. P. (2013). Using eye movements to investigate word frequency effects in children's sentence reading. *School Psychology Review*, 42

Kachergis, G., Shiffrin, R., & Yu, C. (2009). Temporal contiguity in cross-situational statistical learning. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 31, No. 31).

Kempe, V., Brooks, P. J., & Pirott, L. (2001, October). How can child-directed speech facilitate the acquisition of morphology. In *Research on Child Language Acquisition: Proceedings of the 8th conference of the International Association for the Study of Child Language* (pp. 1237-1247). Cascadilla Press.

Klee, T., Carson, D. K., Gavin, W. J., Hall, L., Kent, A., & Reece, S. (1998). Concurrent and predictive validity of an early language screening program. *Journal of Speech, Language, and Hearing Research*, *41*(3), 627-641. https://doi.org/10.1044/jslhr.4103.627

Klin, A. (1991). Young autistic children's listening preferences in regard to speech: A possible characterization of the symptom of social withdrawal. *Journal of Autism and Developmental Disorders*, 21(1), 29–42. https://doi.org/10.1007/bf02206995

Krueger, L. E. (1975). Familiarity effects in visual information processing. *Psychological Bulletin*, 82, 949–974. doi:10.1037/0033-2909.82.6.949 Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. Psychological Review, 99, 22–44. https://doi.org/10.1037/0033-2909.82.6.949

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior research methods*, *44*(4), 978-990. https://doi.org/10.3758/s13428-012-0210-4

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*(3), 299-321. https://doi.org/10.1016/0885-2014(88)90014-7

Larsen, S. F., Schrauf, R. W., Fromholt, P., & Rubin, D. C. (2002). Inner speech and bilingual autobiographical memory: A Polish-Danish cross-cultural study. *Memory*, *10*(1), 45-54. https://doi.org/10.1080/09658210143000218

Law, J., & Roy, P. (2008). Parental report of infant language skills: A review of the development and application of the Communicative Development Inventories. *Child and Adolescent Mental Health*, *13*(4), 198-206. https://doi.org/10.1111/j.1475-3588.2008.00503.x

Lazenby, D. C., Sideridis, G. D., Huntington, N., Prante, M., Dale, P. S., Curtin, S., ... & Akshoomoff, N. (2016). Language differences at 12 months in infants who develop autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*(3), 899-909. https://doi.org/10.1007/s10803-015-2632-1

Lewis, M. M. (1963). *Language, Thought and Personality: In Infancy and Childhood.* George G. Harrap & Co. Ltd.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism diagnostic observation schedule*. Los Angeles, CA: Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism Diagnostic Observation Schedule*, Second Edition (ADOS-2). Torrance, CA: Western Psychological Services.

Luyster, R., Lopez, K., & Lord, C. (2007). Characterizing communicative development in children referred for autism spectrum disorders using the MacArthur-Bates Communicative Development Inventory (CDI). *Journal of Child Language*, 34(3), 623-654. https://doi.org/10.1017/s0305000907008094

Macroy-Higgins, M., & Montemarano, E. A. (2016). Attention and word learning in toddlers who are late talkers. *Journal of Child Language*, *43*(5), 1020. https://doi.org/10.1017/s0305000915000379

MacRoy-Higgins, M., Shafer, V. L., Fahey, K. J., & Kaden, E. R. (2016). Vocabulary of Toddlers Who Are Late Talkers. *Journal of Early Intervention*, *38*(2), 118–129. https://doi.org/10.1177/1053815116637620

MacWhinney, B. J. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.

Marian, V., & Neisser, U. (2000). Language-dependent recall of autobiographical memories. *Journal of Experimental Psychology: General*, *129*(3), 361. https://doi.org/10.1037/0096-3445.129.3.361

Marulis, L. M., & Neuman, S. B. (2010). The effects of vocabulary intervention on young children's word learning: A meta-analysis. *Review of educational research*, *80*(3), 300-335. https://doi.org/10.3102/0034654310377087

Matsumoto, A., & Stanny, C. (2006). Language-dependent access to autobiographical memory in Japanese-English bilinguals and US monolinguals. *Memory*, *14*(3), 378-390. https://doi.org/10.1080/09658210500365763

McCauley, S. M., Monaghan, P., & Christiansen, M. H. (2015). Language Emergence in Development. *The handbook of language emergence*, *87*, 415. https://doi.org/10.1002/9781118346136.ch19

McDonald, S. A., & Shillcock, R. C. (2001). Rethinking the word frequency effect: The neglected role of distributional information in lexical processing. *Language and Speech*, *44*(3), 295-322. https://doi.org/10.1177/00238309010440030101

McGregor, K. K., Newman, R. M., Reilly, R. M., & Capone, N. C. (2002). Semantic representation and naming in children with specific language impairment. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2002/081)

McMurray, B. (2007). Defusing the childhood vocabulary explosion. *Science*, 317(5838), 631-631. https://doi.org/10.1126/science.1144073

Meints, K., & Fletcher, K. L. (2001). Toddler communicative development inventory, a UK adaptation of the MacArthur-Bates communicative development inventories. *University of Lincoln Babylab*.

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, *90*(1), 91-117. https://doi.org/10.1016/s0010-0277(03)00140-9

Mintz, T. H., Newport, E. L., & Bever, T. G. (2002). The distributional structure of grammatical categories in speech to young children. *Cognitive Science*, *26*(4), 393-424. https://doi.org/10.1207/s15516709cog2604_1

Mitchell, S., Brian, J., Zwaigenbaum, L., Roberts, W., Szatmari, P., Smith, I., & Bryson, S. (2006). Early language and communication development of infants later diagnosed with autism spectrum disorder. *Journal of Developmental and Behavioral Pediatrics*, 27, S69–S78. https://doi.org/10.1097/00004703-200604002-00004

Moore, C., Dailey, S., Garrison, H., Amatuni, A., & Bergelson, E. (2019). Point, walk, talk: Links between three early milestones, from observation and parental report. *Developmental Psychology*. https://doi.org/10.1037/dev0000738.supp

Moyle, M. J., Ellis Weismer, S., Evans, J. L., & Lindstrom, M. J. (2007). Longitudinal relationships between lexical and grammatical development in typical and late-talking children, *Journal of Speech, Language, and Hearing Research*, 50(2), 508–528. doi:10.1044/1092-4388(2007/035).

Mueller, V., Sepulveda, A., & Rodriguez, S. (2014). The effects of baby sign training on child development. *Early Child Development and Care*, *184*(8), 1178-1191. https://doi.org/10.1080/03004430.2013.854780

Mullen, E. M. (1995). *Mullen scales of early learning* (pp. 58-64). Circle Pines, MN: AGS.

Müller, O., Duñabeitia, J. A., & Carreiras, M. (2010). Orthographic and associative neighborhood density effects: What is shared, what is different? *Psychophysiology*, *47*(3), 455-466. https://doi.org/10.1111/j.1469-8986.2009.00960.x

Naigles, L. R., Kelty, E., Jaffery, R., & Fein, D. (2011). Abstractness and continuity in the syntactic development of young children with autism. *Autism Research*, *4*(6), 422-437. https://doi.org/10.1002/aur.223

Nelson, K. (1973). Some evidence for the cognitive primacy of categorization and its functional basis. *Merrill-Palmer Quarterly of Behavior and Development*, *19*(1), 21-39.

Nelson, A. B., & Shiffrin, R. M. (2006). Modeling the effects of induced frequency. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 28, No. 28).

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 402-407. https://doi.org/10.3758/bf03195588

Nelson, D. L., Schreiber, T. A., & McEvoy, C. L. (1992). Processing implicit and explicit representations. *Psychological Review*, *99*(2), 322. https://doi.org/10.1037/0033-295x.99.2.322

Nematzadeh, A., Fazly, A., & Stevenson, S. (2014). Structural differences in the semantic networks of simulated word learners. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 36, No. 36).

Newman, R. S., Rowe, M. L., & Ratner, N. B. (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. https://doi.org/10.1017/s0305000915000446

Olswang, L. B., Long, S. H., & Fletcher, P. (1997). Verbs in the emergence of word combinations in young children with specific expressive language impairment.

International Journal of Language & Communication Disorders, *32*(2s), 15-33. https://doi.org/10.1111/j.1460-6984.1997.tb01622.x

Osterling, J. A., Dawson, G., & Munson, J. A. (2002). Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Development and Psychopathology*, 14(2), 239–251. https://doi.org/10.1017/s0954579402002031

Pagán, A., & Nation, K. (2019). Learning words via reading: Contextual diversity, spacing, and retrieval effects in adults. *Cognitive Science*, 43(1), e12705. https://doi.org/10.1111/cogs.12705

Pagán, A., Bird, M., Hsiao, Y., & Nation, K. (2020). Both semantic diversity and frequency influence children's sentence reading. *Scientific Studies of Reading*, 24(4), 356-364. https://doi.org/10.31234/osf.io/8wk4r

Parish-Morris, J. (2011). *Relational vocabulary in preschoolers with autistic spectrum disorder: The role of dynamic spatial concepts and social understanding*. Temple University.

Paul, R. (1991). Profiles of toddlers with slow expressive language development. *Topics in Language Disorders, 11*(4), 1–13. https://doi.org/10.1097/00011363-199111040-00003

Paul, R., & Elwood, T. J. (1991). Maternal linguistic input to toddlers with slow expressive language development. *Journal of Speech, Language, and Hearing Research*, *34*(5), 982-988. https://doi.org/10.1044/jshr.3405.982

Paul, R., & Shiffer, M. E. (1991). Communicative initiations in normal and late-talking toddlers. *Applied Psycholinguistics*, *12*(4), 419-431. https://doi.org/10.1017/s0142716400005853

Payakachat, N., Tilford, J. M., & Ungar, W. J. (2016). National Database for Autism Research (NDAR): Big data opportunities for health services research and health technology assessment. *PharmacoEconomics*, *34*(2), 127-138. https://doi.org/10.1007/s40273-015-0331-6

Petruccelli, N., Bavin, E. L., & Bretherton, L. (2012). Children with specific language impairment and resolved late talkers: Working memory profiles at 5 years. *Journal*

of Speech, Language, and Hearing Research. https://doi.org/10.1044/1092-4388(2012/11-0288)

Pexman, P. M., Hargreaves, I. S., Siakaluk, P. D., Bodner, G. E., & Pope, J. (2008). There are many ways to be rich: Effects of three measures of semantic richness on visual word recognition. *Psychonomic Bulletin & Review*, *15*(1), 161-167. https://doi.org/10.1037/e527342012-696

Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, *36*(4), 329-347. https://doi.org/10.1017/s0140525x12001495

Pierce, K., Conant, D., Hazin, R., Stoner, R., & Desmond, J. (2011). Preference for geometric patterns early in life as a risk factor for autism. *Archives of General Psychiatry*, 68(1), 101–109. https://doi.org/10.1001/archgenpsychiatry.2010.113

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press

Quinn, P. C., & Johnson, M. H. (2000). Global-before-basic object categorization in connectionist networks and 2-month-old infants. *Infancy*, *1*(1), 31-46. https://doi.org/10.1207/s15327078in0101_04

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

Redington, M., Chater, N., & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, *22*(4), 425-469. https://doi.org/10.1207/s15516709cog2204_2

Reilly, S., Wake, M., Bavin, E. L., Prior, M., Williams, J., Bretherton, L., ... & Ukoumunne, O. C. (2007). Predicting language at 2 years of age: a prospective community study. *Pediatrics*, *120*(6), e1441-e1449. https://doi.org/10.1542/peds.2007-0045

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., ... & Bretherton, L. (2010). Predicting language outcomes at 4 years of age: findings from Early

Language in Victoria Study. *Pediatrics*, *126*(6), e1530-e1537. https://doi.org/10.1542/peds.2010-0254

Rescorla, L. (1989). The Language Development Survey: A screening tool for delayed language in toddlers. *Journal of Speech and Hearing Disorders*, *54*(4), 587-599. https://doi.org/10.1044/jshd.5404.587

Rescorla, L. (2002). Language and reading outcomes to age 9 in late-talking toddlers. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2002/028)

Rescorla, L. (2005). Age 13 language and reading outcomes in late-talking toddlers. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2005/031)

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 & Hearing Research*, 52(1), 16. doi.org/10.1044/1092-4388(2008/07-0171)

Rescorla, L. (2011). Late talkers: Do good predictors of outcome exist? *Developmental Disabilities Research Reviews*, *17*(2), 141-150. https://doi.org/10.1002/ddrr.1108

Rescorla, L. A. (1980). Overextension in early language development. *Journal of Child Language*, *7*(2), 321-335. https://doi.org/10.1017/s0305000900002658

Rescorla, L., & Achenbach, T. M. (2002). Use of the Language Development Survey (LDS) in a national probability sample of children 18 to 35 months old. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2002/059)

Rescorla, L., & Fechnay, T. (1996). Mother child synchrony and communicative reciprocity in late-talking toddlers. *Journal of Speech, Language, and Hearing Research*, *39*(1), 200-208. https://doi.org/10.1044/jshr.3901.200

Rescorla, L., & Goossens, M. (1992). Symbolic play development in toddlers with expressive specific language impairment (SLI-E). *Journal of Speech, Language, and Hearing Research*, *35*(6), 1290-1302. https://doi.org/10.1044/jshr.3506.1290

Rescorla, L., & Roberts, J. (2002). Nominal versus verbal morpheme use in late talkers at ages 3 and 4. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2002/098)

Rescorla, L., & Safyer, P. (2013). Lexical composition in children with autism spectrum disorder (ASD). *Journal of Child Language*, 40(1), 47–68. https://doi.org/10.1017/S0305000912000232

Rescorla, L., & Turner, H. L. (2015). Morphology and syntax in late talkers at age 5. *Journal of Speech, Language, and Hearing Research*, *58*(2), 434-444. https://doi.org/10.1044/2015_jslhr-l-14-0042

Rescorla, L., Dahlsgaard, K., & Roberts, J. (2000). Late-talking toddlers: MLU and IPSyn outcomes at 3; 0 and 4; 0. *Journal of Child Language*, *27*(3), 643. https://doi.org/10.1017/s0305000900004232

Rescorla, L., Lee, Y. M. C., Oh, K. J., & Kim, Y. A. (2013). Lexical development in Korean: Vocabulary size, lexical composition, and late talking. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2012/11-0329)

Rescorla, L., Mirak, J., & Singh, L. (2000). Vocabulary growth in late talkers: Lexical development from 2; 0 to 3; 0. *Journal of Child Language*, *27*(2), 293-311. https://doi.org/10.1017/s030500090000413x

Rescorla, L., Roberts, J., & Dahlsgaard, K. (1997). Late talkers at 2: Outcome at age 3. *Journal of Speech, Language, and Hearing Research*, *40*(3), 556-566. https://doi.org/10.1044/jslhr.4003.556

Rice, M. L. (2013). Language growth and genetics of specific language impairment. *International Journal of Speech-Language Pathology*, *15*(3), 223-233. https://doi.org/10.3109/17549507.2013.783113

Rice, M. L. (2003). A unified model of specific and general language delay: Grammatical tense as a clinical marker for unexpected variation. In Y. Levt & J. Schaeffer (Eds.), *Language competence across populations: Toward a definition of specific language impairment* (pp. 63–95). Mahwah, NJ: Erlbaum

Rosa, E., Tapia, J. L., & Perea, M. (2017). Contextual diversity facilitates learning new words in the classroom. *PLoS One,* 12(6), e0179004. https://doi.org/10.1371/journal.pone.0179004

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*(3), 382-439. https://doi.org/10.1016/0010-0285(76)90013-x

Roux, C., Fraïsse, C., Castric, V., Vekemans, X., Pogson, G. H., & Bierne, N. (2014). Can we continue to neglect genomic variation in introgression rates when inferring the history of speciation? A case study in a Mytilus hybrid zone. *Journal of Evolutionary Biology*, 27(8), 1662-1675.

Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development and child vocabulary skill. *Journal of Child Language*. https://doi.org/10.1017/s0305000907008343

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, *83*(5), 1762-1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x

Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences*, *112*(41), 12663-12668. https://doi.org/10.1073/pnas.1419773112

Rutter, M., Le Couteur, A., & Lord, C. (2003). *Autism Diagnostic Interview-Revised*. Los Angeles, CA: Western Psychological Services.

Saffran, J. R., Aslin, R. N. and Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science* 274, 1926–28. https://doi.org/10.1126/science.274.5294.1926

Sahni, S. D., & Rogers, T. T. (2008, August). Sound versus meaning: What matters most in early word learning? In *2008 7th IEEE International Conference on Development and Learning* (pp. 280-285). IEEE.

Sailor, K. M. (2013). Is vocabulary growth influenced by the relations among words in a language learner's vocabulary? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(5), 1657. https://doi.org/10.1037/a0032993

Scarborough, R. A. (2004). *Coarticulation and the structure of the lexicon* (p. 154). University of California, Los Angeles.

Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *CARS: The childhood autism rating scale*. Western Psychological Services.

Schwartz, R. G., & Terrell, B. Y. (1983). The role of input frequency in lexical acquisition. *Journal of Child Language*, *10*(1), 57-64. https://doi.org/10.1017/s0305000900005134

Scott, R. M., & Fisher, C. (2012). 2.5-year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition*, 122(2), 163-180. https://doi.org/10.1016/j.cognition.2011.10.010

Shrestha, R., Shaw, B., Woyczynski, W. A., Thomas, P. J., Fritsch, T., & Lerner, A. J. (2015). Growth and evolution of category fluency network graphs. *Journal of Systems and Integrative Neuroscience*, 1(1), 6-13. https://doi.org/10.15761/jsin.1000103

Shulman, C., & Guberman, A. (2007). Acquisition of verb meaning through syntactic cues: A comparison of children with autism, children with specific language impairment (SLI) and children with typical language development (TLD). *Journal of Child Language*, *34*(2), 411-423. https://doi.org/10.1017/s0305000906007963

Singleton, N. C. (2018). Late talkers: why the wait-and-see approach is outdated. *Pediatric Clinics*, *65*(1), 13-29. https://doi.org/10.1016/j.pcl.2017.08.018

Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, *61*(1-2), 39-91. https://doi.org/10.1016/s0010-0277(96)00728-7

Snyder, L. S., Bates, E., & Bretherton, I. (1981). Content and context in early lexical development. *Journal of Child Language, 8*(3), 565–582. https://doi.org/10.1017/S0305000900003433

Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Stella, M., Beckage, N. M., & Brede, M. (2017). Multiplex lexical networks reveal patterns in early word acquisition in children. *Scientific reports*, *7*, 46730. https://doi.org/10.1038/srep46730

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, *29*(1), 41-78. https://doi.org/10.1207/s15516709cog2901_3

Steyvers, M., Shiffirin, R. M., & Nelson, D. L. (2005). Word association spaces for predicting semantic similarity effects in episodic memory. In A. F. Healy (Ed.), *Experimental cognitive psychology and its applications* (pp. 237–249). Washington, DC: American Psychological Association. https://doi.org/10.1037/10895-018

Stokes, S. F. (2010). Neighborhood density and word frequency predict vocabulary size in toddlers. *Journal of Speech, Language, and Hearing Research.* https://doi.org/10.1044/1092-4388(2009/08-0254)

Stokes, S. F., de Bree, E., Kerkhoff, A., Momenian, M., & Zamuner, T. (2019). Phonology, Semantics, and the Comprehension–Expression Gap in Emerging Lexicons. *Journal of Speech, Language, and Hearing Research,* 62(12), 4509-4522. https://doi.org/10.1044/2019_jslhr-19-00177

Storkel, H. L., Lee, S. Y. (2011). The independent effects of phonotactic probability and neighbourhood density on lexical acquisition by preschool children. *Language and Cognitive Processes*, 26, 191–211. https://doi.org/10.1080/01690961003787609

Stothard, S. E., Snowling, M. J., Bishop, D. V., Chipchase, B. B., & Kaplan, C. A. (1998). Language-impaired preschoolers: A follow-up into adolescence. *Journal of Speech, Language, and Hearing Research*, *41*(2), 407-418. https://doi.org/10.1044/jslhr.4102.407

Swensen, L. D., Kelley, E., Fein, D., & Naigles, L. R. (2007). Processes of language acquisition in children with autism: Evidence from preferential looking. *Child Development*, 78, 542–557. https://doi.org/10.1111/j.1467-8624.2007.01022.x

Tager-Flusberg, H. (1992). Autistic children's talk about psychological states: Deficits in the early acquisition of a theory of mind. *Child Development*, 63(1), 161–172. https://doi.org/10.2307/1130910

Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1430), 303-314. https://doi.org/10.1098/rstb.2002.1198

Taylor, C., Zubrick, S., & Rice, M. L. (2013). Population and public health perspectives on late language emergence at 24 months as a risk indicator for language impairment at 7 years. In *Late talkers: Language development, interventions, and outcomes* (pp. 23-40). Brookes Publishing. https://doi.org/10.1044/1092-4388(2007/106)

Tek, S., Jaffery, G., Fein, D., & Naigles, L. R. (2008). Do children with autism spectrum disorders show a shape bias in word learning? *Autism Research*, 1, 208–222. https://doi.org/10.1002/aur.38

Thal, D. J., & Tobias, S. (1992). Communicative gestures in children with delayed onset of oral expressive vocabulary. *Journal of Speech, Language, and Hearing Research*, *35*(6), 1281-1289. https://doi.org/10.1044/jshr.3506.1289

Thal, D. J., Marchman, V. A., & Tomblin, J. B. (2013). Late-talking toddlers: Characterization and prediction of continued delay. In L. A. Rescorla & P. S. Dale (Eds.), *Communication and language intervention series. Late talkers: Language development, interventions, and outcomes* (p. 169–201). Paul H Brookes Publishing Co.

Thal, D., & Bates, E. (1988). Language and gesture in late talkers. *Journal of Speech, Language, and Hearing Research*, *31*(1), 115-123. https://doi.org/10.1044/jshr.3101.115

Thal, D., Tobias, S., & Morrison, D. (1991). Language and gesture in late talkers: A 1-year follow-up. *Journal of Speech, Language, and Hearing Research*, *34*(3), 604-612. https://doi.org/10.1044/jshr.3403.604

Thomas, M. S., & Knowland, V. C. (2014). Modeling mechanisms of persisting and resolving delay in language development. *Journal of Speech, Language, and Hearing Research,* 57(2), 467-483. https://doi.org/10.1044/2013_jslhr-l-12-0254

Thomas, M. S., Annaz, D., Ansari, D., Scerif, G., Jarrold, C., & Karmiloff-Smith, A. (2009). Using developmental trajectories to understand developmental disorders. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/1092-4388(2009/07-0144)

Tomasello, M., & Mervis, C. B. (1994). The instrument is great, but measuring comprehension is still a problem. *Monographs of the Society for Research in child Development*, *59*(5), 174-179. https://doi.org/10.1111/j.1540-5834.1994.tb00186.x

Unger, L., Vales, C., & Fisher, A. V. (2020). The Role of Co-Occurrence Statistics in Developing Semantic Knowledge. *Cognitive Science*, *44*(9), e12894. https://doi.org/10.31234/osf.io/qmy2n

Venker, C. E., Kover, S. T., & Weismer, S. E. (2016). Brief report: Fast mapping predicts differences in concurrent and later language abilities among children with ASD. *Journal of Autism and Developmental Disorders*, *46*(3), 1118-1123. https://doi.org/10.1007/s10803-015-2644-x

Vergara-Martínez, M., Comesaña, M., & Perea, M. (2017). The ERP signature of the contextual diversity effect in visual word recognition. *Cognitive, Affective, & Behavioral Neuroscience*, *17*(3), 461-474. https://doi.org/10.3758/s13415-016-0491-7

Vigil, D. C., Hodges, J., & Klee, T. (2005). Quantity and quality of parental language input to late-talking toddlers during play. *Child Language Teaching and Therapy*, *21*(2), 107-122. https://doi.org/10.1191/0265659005ct284oa

Vitevitch, M. S. (2019). *Network Science in Cognitive Psychology*. New York: Routledge. https://doi.org/10.4324/9780367853259

Vitevitch, M. S., Chan, K. Y., & Goldstein, R. (2014). Insights into Failed Lexical Retrieval from Network Science. *Cognitive Psychology*, *68*, 1–32. http://doi.org/10.1016/j.cogpsych.2013.10.002

Vitevitch, M. S., Ercal, G., & Adagarla, B. (2011). Simulating Retrieval from a Highly Clustered Network: Implications for Spoken Word Recognition. *Frontiers in Psychology* ,2(December), 1–10. http://doi.org/10.3389/fpsyg.2011.00369

Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology*, *108*(2), 394-401. https://doi.org/10.1016/j.jecp.2010.09.011

Vuksanovic, J., & Bjekic, J. (2013). Developmental relationship between language and joint attention in late talkers. *Research in Developmental Disabilities*, *34*(8), 2360-2368. https://doi.org/10.1016/j.ridd.2013.04.017

Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology*, *36*(5), 571. https://doi.org/10.1037/0012-1649.36.5.571

Weisleder, A., & Waxman, S. R. (2010). What's in the input? Frequent frames in child-directed speech offer distributional cues to grammatical categories in Spanish and English. *Journal of Child Language*, *37*(5), 1089-1108. https://doi.org/10.1017/s0305000909990067

Whitehurst, G. J., Fischel, J. E., Lonigan, C. J., Valdez-Menchaca, M. C., DeBaryshe, B. D., & Caulfield, M. B. (1988). Verbal interaction in families of normal and expressive-language-delayed children. *Developmental Psychology*, *24*(5), 690. https://doi.org/10.1037/0012-1649.24.5.690

Williams, E., Reddy, V., & Costall, A. (2001). Taking a closer look at functional play in children with autism. *Journal of Autism and Developmental Disorders*, *31*(1), 67-77. https://doi.org/10.1023/a:1005665714197

Wood, S. N. (2001). mgcv: GAMs and generalized ridge regression for R. *R news*, 1(2), 20-25.

Wood, S. N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(1), 3-36. https://doi.org/10.1111/j.1467-9868.2010.00749.x

Yurovsky, D., Bion, R. A., Smith, L. B., & Fernald, A. (2012). Mutual exclusivity and vocabulary structure. *N. Miyake, D. Peebles, & RP Cooper (Eds.)*, 1197-1202.

Zeileis, A., & Hothorn, T. (2002). Diagnostic checking in regression relationships. *R News* 2(3), 7-10. URL https://CRAN.R-project.org/doc/Rnews/

Ziatas, K., Durkin, K., & Pratt, C. (1998). Belief term development in children with autism, Asperger syndrome, specific language impairment, and normal development: Links to theory of mind development. *Journal of Child Psychology and Psychiatry*, 39, 755–763. https://doi.org/10.1111/1469-7610.00374

Zubrick, S. R., Taylor, C. L., Rice, M. L., & Slegers, D. W. (2007). Late language emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates. *Journal of Speech, Language, and Hearing Research.* https://doi.org/10.1044/1092-4388(2007/106)

Appendix A

Table A.1 Wilcoxon Rank Sum Test for each Syntactic Class classified using Bates et al.'s approach (1994).

								ASD vs TT		ASD vs LT		TT vs LT	
Syntactic Class	Vocabulary Size	Median ASD	Mean ASD	Median LT	Mean LT	Median TT	Mean TT	*W*	*p*	*W*	*p*	*W*	*p*
Nouns	(0,25]	0.769	0.660	0.854	0.778	0.875	0.813	22526.0	0.058	4909.0	0.160	179629.0	1.130
	(25,50]	0.650	0.654	0.727	0.714	0.800	0.781	4573.5	0.001	585.0	0.299	42366.5	0.001
	(50,75]	0.723	0.688	0.697	0.714	0.786	0.777	2151.5	0.627	329.0	2.980	17375.0	0.001
	(75,100]	0.769	0.769	0.750	0.756	0.768	0.759	728.5	2.815	43.5	2.815	3666.5	2.815
	(100,150]	0.813	0.763	0.743	0.736	0.756	0.753	3636.0	1.038	289.5	1.038	9394.5	1.038
	(150,200]	0.714	0.723	0.714	0.719	0.733	0.730	1606.0	2.581	81.0	2.581	3243.0	2.581
	(200,250]	0.728	0.695	0.696	0.687	0.714	0.709	2088.5	2.723	88.0	2.544	1914.5	2.544
Predicates	(0,25]	0.100	0.214	0.000	0.120	0.000	0.094	35810.5	0.031	7231.5	0.066	171396.0	2.003
	(25,50]	0.231	0.218	0.130	0.141	0.095	0.117	13690.0	0.000	1119.0	0.007	26466.0	0.007
	(50,75]	0.155	0.182	0.177	0.170	0.122	0.133	3550.0	0.718	344.0	2.476	9478.5	0.006
	(75,100]	0.175	0.186	0.162	0.165	0.151	0.157	941.0	1.684	48.0	1.684	3268.5	1.684
	(100,150]	0.138	0.172	0.180	0.189	0.172	0.176	2784.5	1.593	191.5	1.558	7777.5	1.558
	(150,200]	0.222	0.218	0.208	0.200	0.203	0.204	2175.5	1.140	114.0	1.140	3162.5	2.402
	(200,250]	0.241	0.242	0.239	0.236	0.220	0.226	2514.5	1.696	88.0	1.696	1435.0	1.696

Closed class	(0,25]	0.000	0.126	0.000	0.102	0.000	0.093	31531.0	1.393	6511.5	1.393	174445.0	2.786
	(25,50]	0.099	0.128	0.143	0.146	0.087	0.102	10195.0	0.832	728.5	2.143	26018.0	0.010
	(50,75]	0.108	0.130	0.119	0.116	0.082	0.090	3241.0	1.959	302.0	1.959	10183.0	0.055
	(75,100]	0.054	0.045	0.075	0.079	0.077	0.084	359.5	0.801	25.0	1.176	3758.5	2.026
	(100,150]	0.052	0.065	0.068	0.076	0.067	0.071	2493.5	1.029	180.0	1.029	8255.0	2.100
	(150,200]	0.058	0.059	0.074	0.081	0.063	0.066	1602.5	1.692	71.5	1.692	2782.5	1.692
	(200,250]	0.041	0.063	0.071	0.077	0.061	0.065	1640.5	0.443	48.5	0.443	1169.0	0.443

Note. *p*-values were first corrected using the BH method (i.e., corrections accounted for comparisons for the 3 groups of children) then corrected again using the Bonferroni method (i.e., corrections accounted for comparisons for the 3 syntactic classes).

Appendix B

Table B.1 Top-10 Words that Older and Younger Children Differed the Most in Production per Vocabulary Size

Group	(0,25]	(25,50]	(50,75]	(75,100]	(100,150]	(150,200]	(200,250]
Older	mommy	yes	bee	cereal	ice-cream	jacket	lion
	daddy	choo-choo	potty	candy	helicopter	orange (description)	red
	go	please	help	horse	candy	red	jump
	uhoh	go	go-potty	TV	help	shorts	playdough
	ball	ouch	mine	bed	hurt	swing (action)	snow
	ouch	shush	go	drink (food)	TV	blue	black
	car	car	eat	go	pants	plate	chicken (animal)
	moo	hello	please	truck	broken	clean (action)	penguin
	mine	blue	candy	gum	motorcycle	open	stop
	shush	mine	me	hug	bus	ice-cream	broken
Younger	bottle	bottle	bear	bottle	yum	doll	owiebooboo
	hi	book	balloon	cracker	cheese	vroom	pig
	kitty	dog	cracker	button	bear	bunny	toast
	duck	kitty	juice	baa	doll	broom	bib
	balloon	cracker	cheese	out	block	nice	drink (action)
	baby	bird	bottle	bird	bellybutton	toy	call (phone)
	bird	balloon	bath	peas	bath	walk	owl
	yum	cat	airplane	yum	toe	ouch	belt
	that	woof	yum	that	spoon	mine	buttocks/bottom
	book	duck	teddy bear	keys	hat	buttocks/bottom	bellybutton

Note. The above table shows the top-ten words that each group produced more of compared to their size matched equivalent younger or older group.

Appendix C

Procedure for the Estimation of the Best Window Size for Word Co-Occurrence

We tested a total of 25 window sizes on the CHILDES corpus: 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90 and 100 (a description of the corpus used can be found in the Methods section of the manuscript). We computed the contextual diversity values for each word in the Word & Sentences CDI (W&S) in two different ways: the first set of contextual diversity values were extracted from the large matrix of co-occurrence, the *corpus matrix*, where each word in the corpus has a row and a column (19,399 X 19,399); the second set of contextual diversity values were calculated from a reduced version of this first matrix, the *CDI matrix*, in which only the words that appear in the CDI are included (584 X 584, 96 words in the CDI are not in in the corpus). These two matrices were transformed into binary matrices, where values >0 were converted into 1. To compute the contextual diversity values, we added the total sum of the column and the total sum of the row of each word. The values obtained from using the corpus matrix are the *corpus contextual diversity* values, the values obtained from using the CDI matrix are the *CDI contextual diversity* values. The corpus contextual diversity values are higher than the CDI contextual diversity. In total, we created 50 sets of contextual diversity values: 25 window sizes by two matrices (corpus matrix and CDI matrix).

We aimed to find out which of the 50 sets of contextual diversity values predicts AoA the best as well as the order in which these words are produced by children. We used two collections of AoA: Kuperman et al.'s AoA (2012) and our measure of AoA. To compute our AoA, we used the American English CDI W&S item-level data from 5,520 children aged 16 to 30 months downloaded during April 2018 from *Wordbank* (Frank et al., 2017). We carried out a logistic regression on each word to test what age best predict its earliest production. The resulted AoA (*CDI AoA*) is correlated with Kuperman's AoA, r_s = .54, p < .001. To compute the order in which the CDI words are learned by children, we used the same set of CDI data described above. We counted the number of times each word was produced by the children in our samples. We assume that those words that are produced by most children are also

those that are produced earlier. This resulted in a list of words ordered from the most spoken word to the least spoken, each word was assigned an ordinal number.

The CDI words are composed of words with highly semantic content, i.e., nouns and verbs, as well as others with less semantic content, such as prepositions, pronouns, quantifiers and article, among others. Previously, Hills et al.' (2010) utilised Kuperman's AoA to predict the best window size in a linear regression on different types of words, resulting in nouns, verbs and adjectives having the highest R^2s. The finding suggests that children used contextual diversity cues to learn those words with high semantic meaning. For this reason, in our attempt to predict the best set of contextual diversity values, we used two versions of each independent variable (Kuperman's AoA, CDI AoA, Word Order), one with all the CDI words, and another with nouns and verbs only.

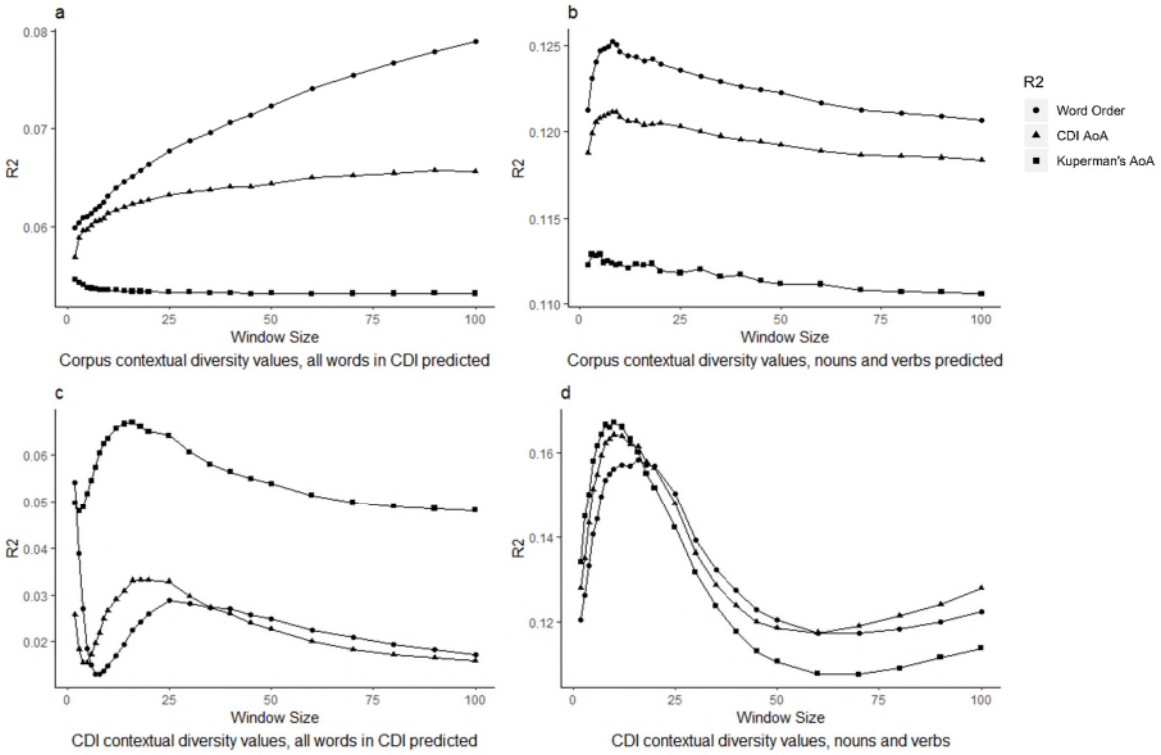
The structure of the four linear regression models performed for each window size is:

1. Corpus Contextual Diversity of all CDI words ~ Word order/CDI AoA/ Kuperman's AoA + Frequency
2. Corpus Contextual Diversity of nouns and verbs ~ Word order/CDI AoA/ Kuperman's AoA + Frequency
3. CDI Contextual Diversity of all CDI words ~ Word order/CDI AoA/Kuperman's AoA + Frequency
4. CDI Contextual Diversity of nouns and verbs ~ Word order/CDI AoA/ Kuperman's AoA + Frequency

The frequency of the words was calculated from the CHILDES corpus. We included word frequency in the equation since frequency is also predictive of AoA and is highly correlated with contextual diversity. Figure C.1 shows the R^2s obtained from each regression. The first observation from the figure is that the trajectories of the R^2s across window sizes from model (d) are similar among the three predictors of interest. R^2s from model (d) are also the highest from the four models tested. From this model, the best window size for each predictor is 16 for Word Order ($R^2 = .16$), and 10 for CDI AoA ($R^2 = .16$) as well as for Kuperman's AoA ($R^2 = .17$). To be consistent with previous research that has used AoA (Hills et al., 2010; Beckage et al., 2011) and since both measures of AoA agreed on the best window size, we decided to utilise a window

size of 10 to compute the contextual diversity values for the words for the current study.

Figure C.1 The R^2 Obtained from each Linear Regression Performed



Appendix D

Correlation Between Order of Word Acquisition and Contextual Diversity

To compute the order in which words are acquired by children, we counted for each word the number of children from Study 1 that had that word in their productive vocabulary. The resulting numbers were a proxy for Kuperman et al.'s AoA (2012), with which has a negative correlation ($r(538) = -.24$, $p < .001$). We computed the contextual diversity of all words in the CHILDES corpus as described in the main text of the manuscript, and then we matched each unique word (n=19,398) with the word order counts, which left most words with a value of 0 (not produced by any child). We found contextual diversity to be positively correlated with word order, $r(19396) = .46$, $p < .001$.

Appendix E

Vocabulary Checklist

a
a lot
about
above
aeroplane / plane
after
all
all gone
alligator / crocodile
am
an
and
angry
animal
ankle
another
ant
any
apple
are
arm
around
asleep
at
aunt
awake
away
baa baa
baby
babysitter
back (Body part)
back (Preposition)
backyard
bad
ball
balloon
banana
bat
bath
bathroom
bathtub / bath
be
beach
beans
bear
because
bed
bedroom
bee
before
behind
belly button / tummy button
belt
bench
beside
better
between
bib
bicycle / bike
big
bin
bird
biscuit
bite
black
blanket
block / brick
blow
blue
boat
book
boot (s)
bottle
bottom
bowl
box
boy
bread
break
breakfast
bring
broken
broom
brother
brown
brush
bubble
bucket
bug
build
bump
bunny / rabbit
bus
but
butter
butterfly
button
buy
by
bye bye
cake
call
camera
camping
can
can / tin
car
careful
carrot
carry
cat
catch
cereal
chair
chalk
chase
cheek
cheese
chewing gum / bubble
chicken (Animal)

chicken (Food)
child
chin
chips
chocolate
chocolate biscuit
choo choo
church / mosque / synagogue /temple
church / mosque / synagogue /temple
circus
clap
clean (Action)
clean (Adjective)
climb
clock
close
cloud
clown
coat
cockadoodledoo
cockerel
coffee
coke
cold
comb
cook
cornflakes
cot
could
countryside
cover
cow
cracker
crayon
crisps
cry
cuddle
cup
cupboard
cut
cute
daddy / dad
dance
dark
day
deer
did /did ya
dining room
dinner
dirty
dish
do
doctor
does
dog
doll
don't
donkey
door
dough
doughnut
down
draw
drawer
dress
drink (Action)
drink (Beverage)
drive
drop
dry (Action)
dry (Adjective)
dryer
duck
dummy
dump
each
ear
eat
egg
elbow
elephant
empty
every
eye
face
fall
farm
fast
feed
find
fine
finger
finish
fire engine
fireman
first
fish (Animal)
fish (Food)
fit
fix
flag
flower
food
foot
for
fork
friend
frog
full
game
garage
garden
gentle
get
giraffe
girl
give
glass
glasses / spec
gloves

go
go potty
gonna / going to
good
goose
gotta / got to
granddad
grandma / nanna /
granny
grapes
grass
green
green beans
grr
hafta / have to
hair
hamburger / burger
hammer
hand
happy
hard
hat
hate
have
he
head
hear
heavy
helicopter
hello
help
hen
her
here
hers
hi
hide
high
high chair
him
his
hit
hold
home
hoover / vacuum
horse
hose-pipe
hot
house
how
hungry
hurry
hurt
i (I)
ice
ice cream
ice lolly
if
inside
into
is
it
jacket
jam
jar
jeans
jelly
juice
jump
jumper / sweater
keys
kick
kiss
kitchen
kitten
knee
knife
knock
know
ladder
lady
lamb
lamp
last
later
lawn mower
leg
lemme / let me
lick
light
like
lion
lips
listen
little
living room / lounge /
lollipop / lolly
long
look
lorry / truck
loud
love
lunch
make
man
me
meat
medicine
melon
meow
milk
mine
mittens
money
monkey
moo
moon
moose
mop
more

morning
motorbike
mouse
mouth
much
muffin / bun
mummy / mum
my
myself
nail
nap / sleep
napkin
nappy
nasty
naughty
neck
necklace
need / need to
new
next to
nice
night
night night
no
noisy
none
nose
not
now
nurse
nuts
of
off
old
on
on top of
open
orange (Colour)
orange (Food)
other
ouch
our
out
outside
oven / cooker
over
paint
pancake
pants / underpants
paper
park
party
pasta / spaghetti
pat-a-cake
path
peanut butter
peas
peekaboo
pen
pencil
penguin
penny
people
person
petrol station
pick
pickle
picnic
picture
pig
pillow
pizza
plant
plate
play
play pen
playground
please
plum
police
pond
pony
poor
pop
popcorn
porch
postman
potato
potty
pour
present
pretend
pretty
pudding
pull
pumpkin
puppy
purse
push
pushchair
put
puzzle
pyjamas
quack
quiet
radio
rain
raincoat
raisin
read
red
refrigerator / fridge
ride
rip
road
rock / stone
rocking chair
roof
room
rubbish

run
sad
salt
same
sandpit
sandwich
sauce
say
scared
scarf
school
scissors
scratch
seaside
see
sellotape
shake
share
she
sheep
shh / hush
shirt
shoe
shop
shopping
shorts
shoulder
shovel / spade
show
shower
shut / close
sick
sing
sink
sister
sit
sky
sleep
sleepy
sleigh
slide (Action)
slide (Object)
slipper
slow
small
smell
smile
snack
snow
snowman
so
soap
sock
sofa / settee
soft
some
soup
spider
spill
splash
spoon
squirrel
stair gate
stairs
stand
star
stay
stick
sticky
stop
story
stove
strawberry
street
stuck
sun
sweep
sweets
swim
swing (Action)
swing (Object)
table
take
talk
taste
tea
teacher
tear
teddy bear
telephone / phone
tell
thank you
that
the
their
them
then
there
these
they
think
thirsty
this
those
throw
tickle
tiger
tights
time
tiny
tired
tissue
to
toast
today
toe
tomato sauce / ketchup
tomorrow
tongue
tonight

too
tooth
toothbrush
touch
towel
town
toy
tractor
train
trainers
tray
tree
tricycle
trousers
try / try to
tummy
tuna
turkey
turn around
turtle
tv / television / telly
uh oh
uncle
under
up
us
vanilla
vitamins
vroom
wait
wake
walk
wanna / want to
was
wash
washer
washing machine /
watch (Action)
watch (Object)
water
we
were
wet
what
wheelbarrow
when
where
which
white
who
why
will
wind
window
windy
wipe
with
wolf
woman
woods
woof
work (Action)
work (Place)
would
write
yellow
yes
yesterday
yoghurt
you
your
yucky
yum
zebra
zip
zoo

Appendix F

Sample

Table F.1 shows the characteristics of the families of the LTs, TTs and LBs. Most mothers had completed some university education (85%; this category includes "Some university", "University degree", "Some postgraduate work", and "Post graduate degree"), whereas only 53% of fathers have attained this level. No differences between the groups were found ($p > .05$, two-sided; Fisher's Exact Test conducted with all levels shown in Table F.1 for Education, excluding "Preferred not to answered" level). Most mothers of TTs and LB were aged 26 to 30 years old, whereas most LTs' mothers were older than 31 years old. Despite this observation, we found no difference between the groups in terms of maternal age ($p > .05$, two-sided); no parental age differences were found between the groups either ($p > .05$, two-sided). With regards to the household income, the vast majority of families earn more than £20,000, with LTs families showing a high proportion of families reaching £65,000 or over. Our statistical analysis showed no income differences between the groups ($p > .05$, two-sided). Most children had no siblings. The same pattern of the number of siblings is shown in every group, with no significant differences ($p > .05$, two-sided). About 60% of the children in the study attended nursery, with the LBs showing the highest percentage and the TTs showing the lowest percentage; no significant differences were found between the groups ($p > .05$, two-sided). Since previous research has shown the positive impact of Baby sign on general development (Mueller, Sepulveda, & Rodrigez, 2014), we also asked parents for the exposure of this language. All groups showed a similar level of exposure to Baby Sign ($p > .05$, two-sided). During the study, they received two children's books every month of participation and a semantic network visualization of their children's vocabularies as they grew throughout the study, one graph per month of participation. The average number of days of participation, i.e., the difference between the first entry of words and the last vocabulary update, was 164.4 (5.3 months), SD = 20.2.

Procedure

The study had a duration of six months. Parents were instructed to download a specially designed application onto their smartphones or tablets for the study. There

were two versions of the app, one for Android devices and another for iOS devices. To start using the app, the parents were given a user name and a password. The app consisted of a searchable vocabulary checklist (described in the previous section) and a voice recorder. Parents were asked at the beginning of the study to mark all the words that their child either "says and understands" or just "understands". They were asked to update this checklist regularly; we made it clear that ideally, we would like them to mark new words as soon as they notice them. Parents had access to the study's website where they were asked to visit to download and read the guide document on how to use the phone application and a guide document on how to proceed to mark words and record audios.

Along with the vocabulary marking, parents were instructed to record audios of themselves using the study app while interacting with their children during different daily routines. They were asked to keep their phones close and not to have any radio or T.V. on, and avoid as much as possible external sources of noise during recordings. Since we sought to capture the natural language surrounding the daily routines, we allowed any caregiver or sibling to appear in the recordings. Every two weeks, participants were sent an email reminding them to submit an audio recording. We specified in the email the routine or topic of the recordings of the fortnight, and some guidance of what type of things could be recorded. In total, they were asked to submit two audio recordings for mealtime, bedtime, bath time, and nappy/potty time, and four audio recordings of playtime. Topics were spread evenly throughout the study so that the time elapsed between any two recordings of the same topic is equal, with exception of playtime which time elapsed between audios were shorter. The topics represent those daily routines that the immense majority of infants and toddlers experience in early life. We asked for a greater number of playtime recordings than other types because the topics covered within play can be more volatile and we sought to capture a wide range of games (e.g., cars, dolls, ball games, cooking.). All audio files were transcribed by a professional UK-based transcription company. Due to the high cost of transcribing all the audio data from these families, we selected only two audios of playtime per family for transcription. All the other audio files were transcribed. A total of 497 audio files were transcribed. The app sent the words updates to an AWS server, and the audio files were sent to a secure server managed by the University of Warwick. When parents had problems submitting audio files through the app, they were

instructed to submit them manually to the study website. Parents that took more than two weeks to submit the required audio, were reminded through email to do so.

Table F.1 Description of the Participating Families

	Typical talkers' families		Late talkers' families		Late bloomers' families		Full sample	
	n	%	n	%	n	%	n	%
Education (Mother/Father)								
Secondary school	1 / 0	5.3 / 0.0	0 / 2	0.0 / 9.1	0 / 0	0.0 / 0.0	1 / 2	1.6 /3.2
Sixth form or college	4 / 4	21.1 / 21.1	0 / 1	0.0 / 4.5	1 / 4	4.5 / 18.2	5 / 9	7.9 / 14.3
Trade/technical/vocational/apprenticeship training	0 / 3	0.0 / 15.8	2 / 6	9.1 / 27.3	1 / 8	4.5 / 36.4	3 /17	4.8 / 27.0
Some university	3 / 1	15.8 / 5.3	2 / 3	9.1 / 13.6	3 / 1	13.6 / 4.5	8 / 5	12.7 / 7.9
University degree	8 / 9	42.1 / 47.4	7 / 6	31.8 / 27.3	9 / 5	40.9 / 22.7	24 / 20	38.1 / 31.7
Some postgraduate work	1 / 0	5.3 / 0.0	3 / 1	13.6 / 4.5	1 / 1	4.5 / 4.5	5 / 2	7.9 / 3.2
Post graduate degree	2 / 2	10.5 / 10.5	7 / 2	31.8 / 9.1	6 / 3	27.3 / 13.6	15 / 7	23.8 / 11.1
Preferred not to answer	0 / 0	0.0 / 0.0	1 / 1	4.5 / 4.5	1 / 0	4.5 / 0.0	2 / 1	3.2 / 1.6
Age (Mother/Father)								
21-25 year old	1 / 2	5.3 / 10.5	2 / 1	9.1 / 4.5	1 / 1	4.5 / 4.5	4 / 4	6.3 / 6.3
26-30 years old	9 / 4	47.4 / 21.1	5 / 3	22.7 / 13.6	9 / 4	40.9 / 18.2	23 / 11	36.5 / 17.5
31-35 years old	6 / 9	31.6 / 47.4	7 / 8	31.8 / 36.4	5 / 8	22.7 / 36.4	18 / 25	28.6 / 39.7
36+ years old	3 / 4	15.8 / 21.1	8 / 10	36.4 / 45.5	7 / 9	31.8 / 40.9	18 / 23	28.6 / 36.5
Household Income per year								

Less than £20,000	2	10.5	1	4.5	1	4.5	4	6.3
£20,000 to £45,000	7	36.8	6	27.3	4	18.2	17	27.0
45,000 to £65,000	7	36.8	6	27.3	9	40.9	22	34.9
More than £65,000	3	15.8	8	36.4	6	27.3	17	27.0
Preferred not to answer	0	0	1	4.5	2	9.1	3	4.8
Siblings								
None	11	57.9	14	63.3	15	68.2	40	63.5
1 sibling	6	31.6	6	27.3	5	22.7	17	27.0
2 siblings	2	10.5	1	4.5	2	9.1	5	7.9
3 or more	0	0.0	1	4.5	0	0.0	1	1.6
Attendance to nursery*	8	42.1	13	59.1	17	77.3	38	60.3
Exposure to Baby sign*	7	36.8	8	36.4	9	40.9	24	38.1

Abbreviations

ABC: Approximate Bayesian Computation.

AIC: Akaike information criterion.

AoA: Age of acquisition.

ASD: Autism Spectrum Disorder.

BH: Benjamini-Hochberg procedure.

BIC: Bayesian information criterion.

CDI: (MacArthur) Communicative Developmental Inventory.

CHILDES: Child Language Data Exchange System (database)

DLD: Developmental Language Disorder

LB: Late bloomer.

LT: Late talker.

MSEL: Mullen Scales of Early Learning.

PA: Preferential Acquisition (model).

PLA: Progressive Lure of Associates (model).

PPA: Progressive Preferential Acquisition (model).

TT: Typical talker.