Differences in the Semantic Structure of the Speech Experienced by Late Talkers, Late Bloomers and Typical Talkers

First submission on 6th of March 2021

Second submission on 5th of August 2021

Third submission on 5th of June 2022

Abstract

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 audio recorded daily routines and marked the new words their child produced over a period of six months. To investigate how language environments differed between talker types and how environments corresponded with children's developing lexicons, we evaluated contextual diversity –a word property that measures semantic richness– and network properties of language environments in tandem with developing vocabularies. The language environment experienced by the three talker types differed in their structural properties, with LT environments being least contextually diverse and least well-connected in relation to network properties. Notably, LBs' language environments were more like those of TTs. Network properties of language environments also correlate with the rate of vocabulary growth over the study period. By comparing differences between language environments and lexical network development, we also observe results consistent with contributions to lexical development from different learning strategies for expressive vocabularies and different environments for receptive vocabularies. We discuss the potential consequences that structural differences in parental speech might have on language development, and the contribution of this work to the debate on quantity versus quality.

Keywords: language environment, early language delay, semantic networks, word acquisition

Introduction

The semantic structure of natural speech has been shown to predict child language acquisition as well as provide a way in which adults may organize their lexicons for efficient language processing (e.g., children: Hills, Maouene, Maouene, Riordan & Smith, 2010; adults: Steyvers & Tenenbaum, 2005; Steyvers, Shiffrin, & Nelson, 2005). What is less well understood is how individual differences in the language environment might be associated with individual differences in language acquisition. As a first step, the current research focuses on the relationship between the structure of the developing lexicons of three groups of children with different early language outcomes (i.e., typical talkers, late talkers, and late bloomers) and the structure of their aggregated language environments.

Two prominent features make ideal targets for quantitative comparisons of language environments and lexical development: contextual diversity and the network properties of language. Contextual diversity is associated with the number of linguistic contexts in which a word appears, and research has shown that words with higher contextual diversity are learned more quickly (e.g., children: Rosa, Tapia, & Perea, 2017; Hills et al., 2010; adults: Pagán & Nation, 2019). Contextual diversity is proposed to aid the enrichment of a words' semantic 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 (i.e., low in contextual diversity) might be associated with low rates of lexical development in early childhood (i.e., early language delay). However, one could expect the opposite given the evidence that consistency in the language input can promote early word acquisition (Roy et al., 2015; Schwab & Lew-Williams, 2016; Rowe, 2012).

Networks can also be derived from the way in which words co-occur with one another. Children learn the semantic relatedness between words from the language they hear, e.g., that *bottle* is associated with *milk*. If one links the words that a child knows using these co-occurrence relations, the results would be a lexical network from which network properties can be evaluated. Network structures can be used to predict vocabulary development and are also informative with respect to individual differences (Bilson, Yoshida, Tran, Woods, & Hills, 2015; Hills, Maouene, Maouene, Sheya & Smith, 2009; Beckage & Colunga, 2019).

Child-directed speech naturally lends itself to the structural analyses of contextual diversity and network properties, and in the present work, we investigate these in relation to the developing vocabularies of 63 children and their language environments. The sample consisted of 19 children with typical language skills (i.e., typical talkers), 22 children that started the study with a language delay and remained delayed by the end of the study (i.e., late talkers), and 22 children that also started the study with a language delay, but who overcame their delay by the end of the study (i.e., late bloomers). 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 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 that contextual diversity facilitates language development, followed by a brief introduction to semantic networks and a description of what is known about the late-talking population in relation to word acquisition.

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 contexts (contextual diversity), the word-referent remains constant across situations, aiding the infant to solve the problem of indeterminacy (Quine, 1960; Hills, 2013). In addition, 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 (i.e., one word only refers to one referent; Markman, 1991). 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 typical age of acquisition, with more contextually diverse words having a higher likelihood of being learned earlier in vocabulary development (e.g., Hills, Maouene, Riordan & Smith, 2010; though see: Roy et al., 2015; Schwab & Lew-Williams, 2016; Rowe, 2012). More recent work using semantic vector space models has confirmed a strong correlation between word connectivity in child-directed speech and rates of word production; highly connected words (e.g., words with high contextual diversity) are produced earlier by toddlers (Amatuni & Bergelson, 2017). Various studies have shown that words learned in many different contexts are easier to process and recognized faster (adults: Goldinger & Azuma, 2004; McDonald & Shillcock, 2001; Nelson & Shiffrin, 2006; Pexman et al., 2008; children: 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 in adults (Adelman, Brown & Quesada, 2006; Johns, Dye & Jones 2016). Contextual diversity can be seen as a word feature, i.e., some words are semantically more versatile and can accompany more different words in speech than others, especially polysemous words (e.g., bank). At the same time, the degree of the contextual diversity of individual words can vary from one language environment to another. In other words, the number of opportunities to learn semantic knowledge from word co-occurrence might be reduced in the speech that some families direct to their children, which might negatively influence word acquisition.

Evidence supporting the role of contextual diversity in word acquisition assumes that the characteristics of the linguistic input are static for the first years of the child's life. However, previous work by Hills (2013) has found contextual diversity of the speech directed to younger children to be lower than that directed to older children. Therefore, it might be the case that a certain level of consistency is good at early stages (perhaps to help children solve the word-referent mapping thanks to repetition of the label with its referent). Later on, parents make their speech more contextually diverse to help their children discover the semantic characteristics associated with each word. In this manuscript, we evaluate three aspects of contextual diversity: 1) whether children with different vocabulary growth rates (i.e., typical talkers, late talkers, and late bloomers) experience different degrees of contextual diversity in the language they hear at home, 2) how contextual diversity changes in child-directed speech as children grow, and 3) if these changes in child-directed speech are similar across the three talker types (LT, LB, and TT).

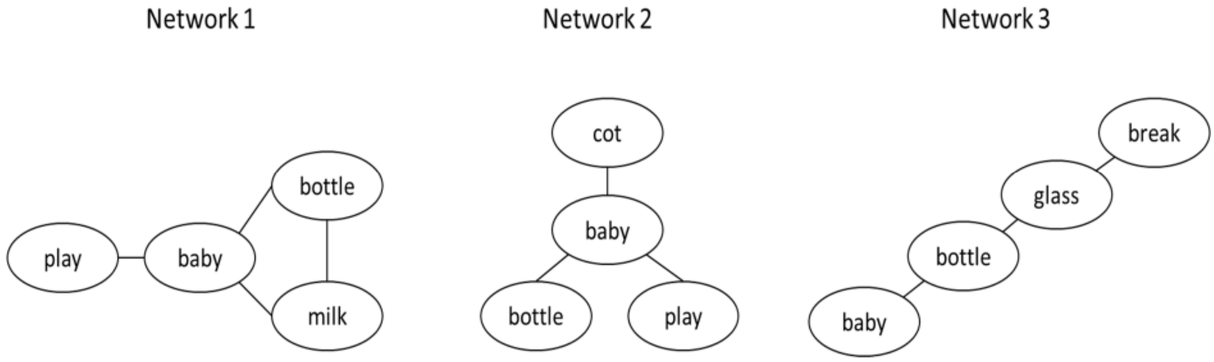
Semantic Networks and Word Acquisition

When a word frequently co-occurs with other words in speech, we create a semantic link between the two in our mind (e.g., *salt* and *pepper*). Words distributed within speech create a complex network of relationships, in which the nodes represent words and the edges indicate a semantic connection between words (see Figure 1 and 2 for an example). This linguistic structure can be quantitatively evaluated using network analysis (see Figure 1 for examples of key measures discussed in this work). As seen in Figure 1, two words can be strongly semantically related if they are separated by one edge (like *baby* and *bottle*), or they can be weakly semantically related if they are linked by a few intermediate nodes (like *baby* and *break*). At the same time, a word can be part of a solid semantic cluster, such as the cluster *baby-bottle-milk* seen in Network 1. Words that are higher in contextual diversity have more semantic links with other words. The early acquisition of high-contextual-diversity words generally leads to well-connected vocabularies.

Previous research showed that children are sensitive to the structure of the language learning environment. Children acquire words in relation to how well-connected they are to other words in the language input rather than how well connected they are to well-connected words in the child's lexicon (Hills et al., 2009; Amatuni & Bergelson, 2017). The structural properties of human language are thought to influence online language processing (e.g., contextual diversity/mean degree: Goldinger & Azuma, 2004; McDonald & Shillcock, 2001; Nelson & Shiffrin, 2006; Pexman et al., 2008; Frances, Martin, & Duñabeitia, 2020; clustering coefficient in phonology and phonological neighborhood density: Vitevitch, Ercal, & Adagarla, 2011; Vitevitch, Chan, & Goldstein, 2014). However, though some network properties are predictive of child language acquisition, it is not yet known how these properties promote word learning in early childhood. The first step in understanding how the environmental structural properties might aid word acquisition is to see how the network properties of the language environment experienced by children with different degrees of language ability correspond to their developing lexicons (see Figure 2 for an illustration of the relation between the structure of a language environment and a child’s vocabulary).

Figure 1

Properties of three same-size networks



Local network properties of node ***baby***

	Net 1	Net 2	Net 3
Degree	3	3	1
Local clustering coefficient	1/3	0	0
Shortest distance ***baby→bottle***	1	1	1
Shortest distance ***baby→break***	-	-	4

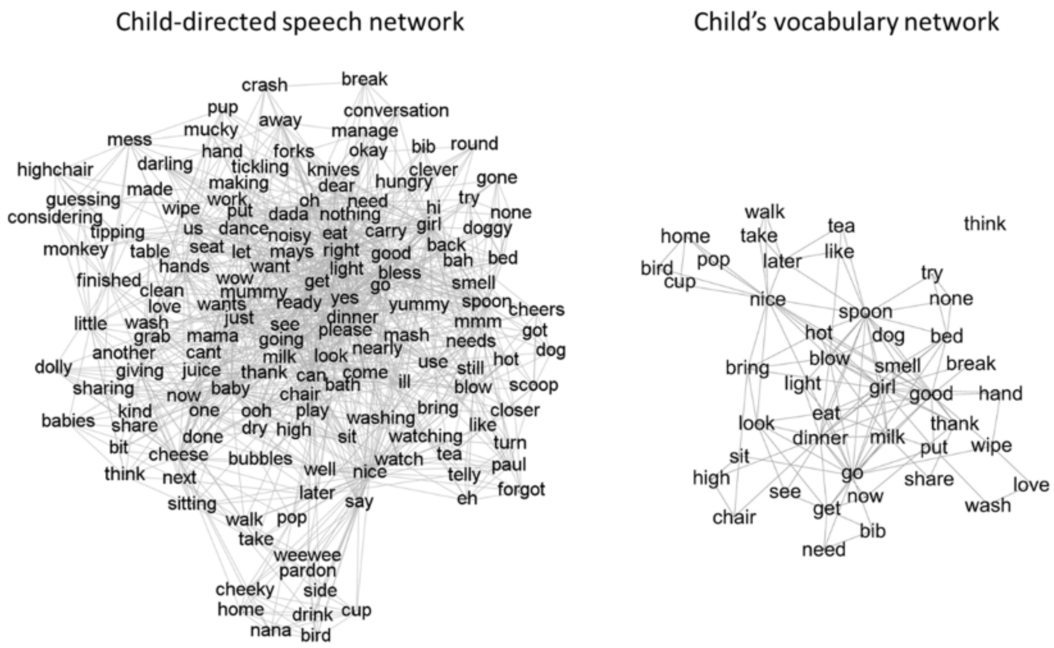
Global network properties

	Net 1	Net 2	Net 3
Mean degree	2	1.5	1.5
Average clustering coefficient	0.7	0	0
Average path length	1.3	1.3	1.6

Note. The local network properties of node *baby* in each network are displayed in the left-hand table. The *degree* of a node is the number of links that that node has with other nodes. *Local clustering coefficient* is a measure of the connectivity between a node's nearest neighbors. It is calculated as the proportion of links between a node's neighbors divided by the maximum number of links that could exist between those neighbors. For example, for Network 1 there are three neighbors (*play*, *bottle*,and *milk*), allowing for three possible connections. But only one connection is observed (*bottle* and *milk*), meaning the local clustering coefficient is 1/3. Shortest distance is the shortest path between each pair of nodes counted in terms of the number of edges. For example, there are three edges between *baby* and *break* in network 3. The global properties of each network are displayed in the right-hand table. To compute these, the local properties are averaged for all the nodes in the network. Well-connected vocabularies are characterized by high mean degree, high clustering coefficient, and low average path length.

Figure 2

The semantic networks of a sample of child-directed speech and the vocabulary of the child exposed to that speech.



	Number of words	Mean degree	Clustering coefficient	Average path length
Child-directed network	151	15.8	0.62	2.1
Child's vocabulary network	44	5.8	0.61	2.3
Random word learner	44	4.4	0.58	2.6

Note. Left: semantic network created from a twenty-minute sample of child-directed speech recorded by a family from the current study. Right: words from the sampled speech that are in the vocabulary checklist used in our study, representing the potential semantic network of the vocabulary of the child exposed to this speech stream. The semantic links shown in the child's network are inherited from her environment. This means that if the child produced all the words in the language input, the two networks (child-directed speech and vocabulary) would have the same structural properties. However, as noted in the main text, children tend to learn more well-connected words earlier. This can be noticed when the child's vocabulary is compared to that of a random learner (bottom row). Here, one-hundred random acquisition networks were computed (i.e., the same number of words were randomly selected from the language environment and connected based on the language environment, following Beckage et al., 2011). Compared to the child's vocabulary, the network of a random learner shows fewer links (mean degree), lower degree of semantic clusters (clustering coefficient), and a larger distance between any pair of nodes (average path length).

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 (e.g., Jiménez, Haebig, & Hills, 2020; Ellis Weismer, 2007). 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. 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. Given the evidence that shows that late-talking toddlers acquire different types of words (i.e., in a different order), further investigations are warranted to determine why this is the case.

Beckage and colleagues (2011) showed that the semantic structure of late-talking toddlers' vocabularies is less clustered and generally less well-connected compared to vocabulary-matched typical talkers, consistent with an impoverished sensitivity to contextual diversity. These results contrast with subsequent work carried out with a much larger sample of children (Jiménez and Hills, 2017), in which late talkers instead exhibited better-connected vocabulary than vocabulary-matched typical talkers. A major assumption in these two studies 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 samples. However, differences in the language structure of individual families are possible, so any conclusions about children's lexical development need to consider the children's unique linguistic environments. Moreover, both studies utilize cross-sectional data. One inconvenience of cross-sectional data in studying late talkers is that some late talkers (late bloomers) catch up with their same-age peers. The present study addresses these methodological challenges by utilizing longitudinal vocabulary data alongside the child-directed speech experienced by each individual child (although aggregated based on children's language ability), as well as a sample containing typical talkers, late talkers, and late bloomers.

Differences in the Language Environment

Many studies have investigated what differentiates the parental speech of late talkers from that of typical talkers. Most research 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 (e.g., 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, differences 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 modeling 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 at early stages, something shown to promote word acquisition in typically talking children (Roy et al., 2015). At the same time, children benefit from 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 linguistic 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.

Current study

The current study aims to 1) investigate the change of the semantic structure of child-directed speech over time and its correlation with children's vocabulary networks, 2) to examine whether children with different rates of lexical acquisition (i.e., typical talkers, late talkers, and late bloomers) experience structurally-different linguistic inputs at home, and 3) to identify similarities and differences in the semantic structure of the vocabulary of typical talkers, late talkers, and late bloomers. 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 research questions are as follows:

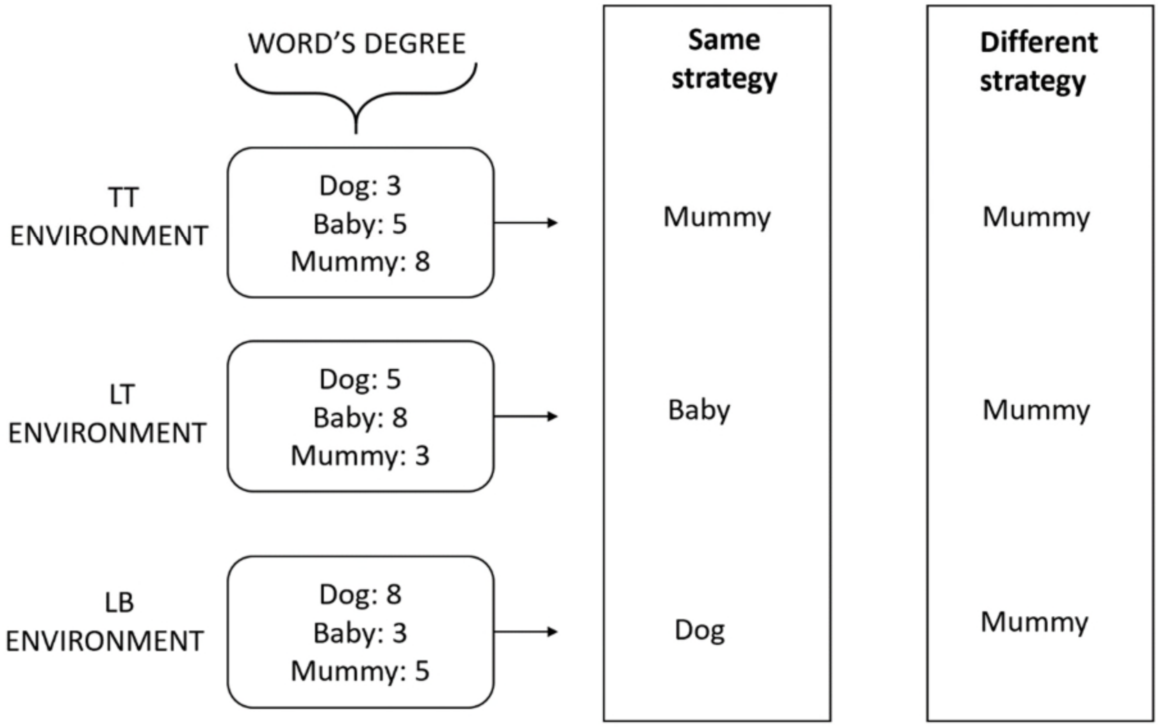
1. **Are there differences in the contextual diversity or network properties between the speech that LTs, TTs, and LBs receive at home?** Given the important role that contextual diversity has been shown to have on word learning, we hypothesized that the speech experienced by children with language delay (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. We expect a similar outcome to Hills (2013)—with younger children receiving less contextually diverse input than older children—but 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 late talkers will still see less contextually-diverse language environments than younger typical talkers who know more words. Our analyses on the network properties of clustering coefficient and average path length are, at this point, exploratory since no study to date has identified structural attributes of speech that might correspond to slow lexical growth. However, as we predict environmental differences based on contextual diversity, we also expect that structural features will differ between our corpora.

2. **Are the network properties of child-directed speech correlated with the network properties of children's vocabularies?** Similarly, do the network properties in a child's vocabulary display the same changes over time as the speech that children of the same age and same vocabulary size hear? If children's vocabularies reflect the language they hear, we should find correlations between the semantic structure of language environments and vocabulary networks.

3. **Do properties of the language environment predict vocabulary growth?** Given the proposed influence of different properties of the language environment, we also ask if these properties are correlated with language change over the duration of our study.
4. **What are the contributions of strategy and environment on lexical development?** Our unique dataset provides a first step towards teasing apart contributions of strategy (do different talker-types use different word learning strategies?) and environment (do different environments bias developmental trajectories?). To answer these questions, we conducted a two-step analysis based on two different ways of determining the semantic edges that link the words within children's vocabularies. In both steps, an edge between two words is formed when they co-occur in speech. Step 1 investigates differences between the lexical networks of LTs, TTs, and LBs using a shared rule for edge formation based on aggregating child-directed speech across talker types. Step 2 investigates differences between the lexical networks of LTs, TTs, and LBs using rules for edge formation based on their different talker-type environments. By examining the differences between the outcomes of these two steps, we can make inferences about whether the developmental trajectories are based on strategic or environmental factors. Specifically, if the different talker-types show different networks based on aggregated edge rules, but similar networks based on their individual talker-type edge rules, then we can infer they are using a similar learning strategy (e.g., learning higher degree words earlier) in their idiosyncratic learning environment. This is, because their environments are different, they learn different words earlier and thus look different when compared with the shared edge rules. Alternatively, if the talker-types show similar networks based on aggregated edge rules, but different networks based on their individual talker-type edge rules, we can infer a difference in strategy—they learn structurally different words in their individual talker-type environments, implying a different learning rule. In Figure 3, we show an illustration of the two types of inferences. We apply this analysis to both receptive and expressive lexical networks.

Figure 3

An Example of the Two Types of Inferences that Can Be Made from Examining the Children's Networks



Note. Each talker-type environment contains the same words, but these words have different network degrees. If LTs, TTs, and LBs were to learn words using a strategy in which the word with the highest degree is learned first (i.e., a preferential acquisition strategy, Hills et al., 2009), the children would learn the highest one from their respective environments (i.e., same strategy). This also means that the differences in their environments contribute to their lexical acquisition. In contrast, if each talker group uses a different strategy (e.g., LTs learning the words with the lowest degree earlier), the three groups might acquire the same word, but the degree of this word differs in their respective environments.

Method

Participants

Ethical permission was granted by [blinded for review] to conduct the current research. The families were recruited (with specific emphasis on recruiting late-talkers) through the child-laboratory database of the [blinded for review] '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 who 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., 1993) 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 Supplemental 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 (1993) 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., 1993) 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; Jiménez & Hills, 2017). 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 productive vocabulary (127 words) was well under the 5th percentile for a 30-month-old child, and by extrapolation, we labeled 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).

Table 1 shows the children's characteristics in each group (see Online Supplemental Materials for histograms of these data). We conducted a series of simple linear regressions 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 2 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 "Postgraduate 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 2 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 language on general development (Mueller, Sepulveda, & Rodriguez, 2014), we also asked parents for the amount of exposure to this language. All groups showed a similar level of exposure to baby sign language ($p > .05$, two-sided). 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), $SD = 20.2$.

Table 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 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 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 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

Table 2

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/ 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
	Postgraduate 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

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 their child either "says and understands" or just "understands". There were no differences between LTs and TTs in terms of vocabulary updates (see Table 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 Supplemental Materials.

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).

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., 2010), a window of size five was moved through 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.

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 inquiry, either the whole corpus or a subsample of the corpus was utilized (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 the 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: 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 neighbors 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).

Corpus Analysis and Semantic Network Analysis

Research question #1: Are there differences in the contextual diversity or network properties between the speech that LTs, TTs, and LBs receive at home?

Each document (transcription) was tagged 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-type 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).

Research question #2: Are the network properties of child-directed speech correlated with the network properties of children's vocabularies?

To explore the correlation between the structural properties of age- and vocabulary-size-matched child-directed speech and the children's networks, we split our corpus into four corpora in two different ways: by the vocabulary size and by the age of the child to whom the speech of the transcription was directed to (vocabulary bins: [1,100], [101-200], [201-300], [301-400]; age bins in months: [13,17], [17.1-21], [21.1-25], [25.1-31]). Utilizing the bootstrapping approach described above (to control for the number of words in each bin), we sampled 1000 times from each of these sub-corpora bins, computed the network properties each time, and then averaged these for each bin. We matched these environmental measures to the data points of each child in our data set. This matching was based on the age and the vocabulary size of the child at that point. For example, a child whose vocabulary and age at a certain point of the study was 50 words and 16 months old was matched to environment network statistics from the vocabulary bin [1,100] and age bin [13,17].

Despite collecting child-directed speech samples from each individual family, the amount collected is not large enough to consider individual linguistic environments in our analysis. There are two main reasons why this is not possible. First, as with our aggregated data analysis above, we would need to control for the number of words produced in each environment, which forces us to employ a bootstrapping approach. And second, we would need to include in our analysis only the words that are common across families within each sampling period. This procedure considerably reduces the amount of data to analyze, which significantly influences the power of our analysis. Similarly, to examine how the network properties in parental speech of individual families change over time, the number of words in each transcript needs to be controlled for, and only common words across transcripts can be evaluated.

Research question #3: Do properties of the language environment predict vocabulary growth?

To calculate the vocabulary growth for each child, we divided the number of words produced by the child during the study by the amount of time they spent participating (i.e., age at the start subtracted from the age at the end of the study). Each child was assigned with the environmental network measures from their talker-type environment aggregated across talkers of the same type. This was necessary due to data limitations associated with predicting children from their own individual environments.

Research question #4: What are the contributions of strategy and environment on lexical development?

To examine whether potential differences in the network properties of the expressive (production) and receptive (comprehension) vocabularies of TTs, LTs, and LBs are associated with strategic or environmental factors, we examined the differences between the outcomes of a two-step analysis. In Step 1, we constructed the children's networks using the same adjacency matrix computed from the whole corpus. In Step 2, we 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 in research question 1 (see above) 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 utilized generalized additive models. The *gam* function in the *mgcv* package in R was used (see Wood 2001). A generalized additive model (GAM) was 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 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.

To compute posthoc power analysis, we used a simulation-based approach because of the complexity given by random effects in the models. We adapted our GAM models into generalized linear mixed models and used the *SIMR* R package to run our simulations (Green & MacLeod, 2016). Across our models, we obtained a power for the predictor 'type talker' between 94% and 97% given our sample size, which is higher than the standard 80% for adequacy (*N*= 64; test: likelihood ratio; *N* simulations= 1000; alpha= 0.05). This study was not preregistered. The code, transcripts and vocabularies used in this study can be found in OSF [LINK-blinded-]. Our transcripts and vocabularies are also available in specific public repositories (Transcripts: CHILDES (https://childes.talkbank.org/); vocabularies in Wordbank (http://wordbank.stanford.edu/) .

Results

Are there differences in the contextual diversity or network properties between the speech that LTs TTs and LBs receive at home?

We examined whether parental speech varies in contextual diversity or in semantic network properties depending on whether it was directed to LTs, TTs, or LBs. Figure 4 shows the contextual diversity averages and network measures for each sample drawn from the LT corpus, TT corpus, or LB corpus, where each dot represents a sampled set of documents (there is a total of 1000 samples/dots for each talker type). The left panels display the relation between each contextual measure and age. The right panels show the relation between each contextual measure and vocabulary size.

Given that we are considering all words in the corpus (not just words in the CDI) to calculate statistics of the environment, contextual diversity and average degree become practically the same measure in this analysis (ρ= .99). This can be seen in Table 3, in which 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.

When age is considered as the x-axis in Figure 4, the general pattern (larger regression line) indicates that contextual diversity and mean degree in child-directed speech decreases as children age, yet the regression lines fitted for each talker group indicate the opposite. This is a clear example of Simpson's paradox, for which group trends violate the aggregate trend. In contrast, when vocabulary size is considered as the x-axis, the general and group regression lines agree: contextual diversity of caregivers' speech is higher for children who produce more words. On the x-marginal density plots, the differences in age and vocabulary size between the groups reflect the talker group average differences described in the Methods (Age: LT>LB>TT; vocabulary size: TT>LB>LT). The y-marginal density plots suggest that the contextual diversity of the speech directed to LTs is lower than that of TTs and LBs. As for TTs and LBs, the contextual diversity of their language environments are more similar.

As seen for contextual diversity and mean degree, the general regression line for clustering coefficient disagrees with the groups' lines, indicating an age and vocabulary size influence on our dependent variable. The y-marginal density plot suggests that the LB group has the highest clustering coefficient, followed by the LT group, with the TT groups displaying the lowest values. With regards to average path length, the potential effect of age and vocabulary on this measure can be deduced by the disagreement between the regression lines. The LT group shows higher average path length than the TT and LB groups, with these latter two groups showing similar averages in this network property.

Table 4 displays the results from the best regression models selected after stepwise comparisons (Bayesian information criterion —BIC— was used as a criterion for model selection). 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. We also explored whether the network properties of child-directed speech change over time in the same way for all groups. To do this, we added interactions between age/vocabulary size and talkers type environment; however, none of the models significantly decreased their BICs after adding the interaction. Standardized regression coefficients for all four models are plotted in Figure 5. Contextual diversity and average degree generally increase as children age, consistent with Hills' findings (2013).

The observation that the language environment of LTs is lower in contextual diversity and average degree than the TTs' and LBs' environments might be due to contextual diversity being conflated with frequency. Further, the different talker environments may vary in the number of words spoken within a particular temporal window. This might bias our results in that audio samples with more silent periods could contain a larger number of topics (since there is a higher chance to switch topics), which could lead to an increase of semantic density and therefore, an increase in contextual diversity. When dividing the total duration of each audio by the number of words that it contains, we found that TT environments present higher semantic density than LT and LB environments (TT vs LT: b= 0.085, CI= [0.05, 0.12], SE= 0.02, p< .001; TT vs LB: b= 0.07, CI= [0.032, 0.11], SE=0.019, p< .001), and LB and LT environments showed similar semantic density (LB vs LT: b= 0.015, CI= [-0.021, 0.051], SE= 0.018, p> .05), although talker type explains little variability in the model, $F(2, 494)$= 10.54; Adj.R2= .04, p< .001. One way to control for frequency is to repeat our bootstrapping procedure and divide each word's contextual diversity value by its frequency in each sample. To control for semantic density, we can take the total contextual diversity value calculated for each sample and divide it by the total length of the audios of the transcripts sampled. We proceeded to repeat our analysis but this time controlling for frequency and semantic density as described above. As can be observed in Figure 6, LTs were still the group with the lowest contextual diversity in their language environments (TT vs LT: b= 688.81, CI= [650.89, 726.74], SE= 19.34, p< .001; LB vs LT: b= 348.71, CI= [326.72, 370.7130], SE=11.22, p< .001). The results also revealed that LBs' environments have lower contextual diversity that the TTs' environments (b= -340.09, CI= [-358.02, -322.17], SE=9.14, p< .001; F(3, 2996)= 5306; Adj.R2= .84, p< .001). Equally as before, the inclusion of age increased the predictability power of the model, with contextual diversity increasing as children age (vocabulary was not a significant predictor when age was included in the model, b= 96.95, CI= [81.43, 112.47], SE=7.91, p< .001).

Clustering coefficient decreases in caregiver speech as children age; it also generally decreases in lexical networks as vocabularies grow[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.

[1] After checking Variance Inflation Factors we found age and vocabulary to be highly positively correlated. 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.

Table 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

Table 4

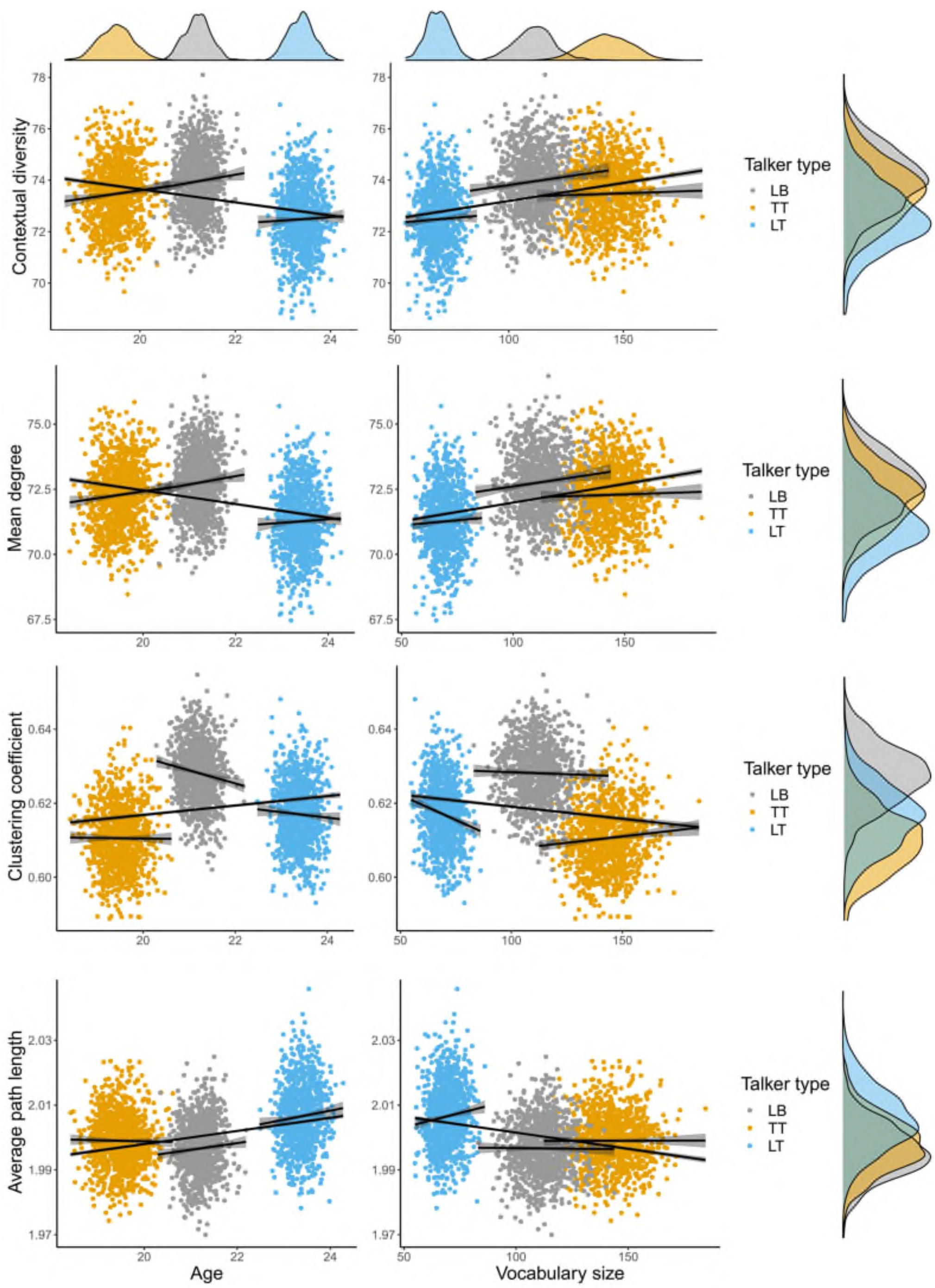
Regression Analysis Predicting Contextual Diversity and Network Properties in Child-Directed Speech

Dependent Variable	Independent Variables	b	CI	SE	p	F (df)	Adj. R^2
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 standardized.

Figure 4

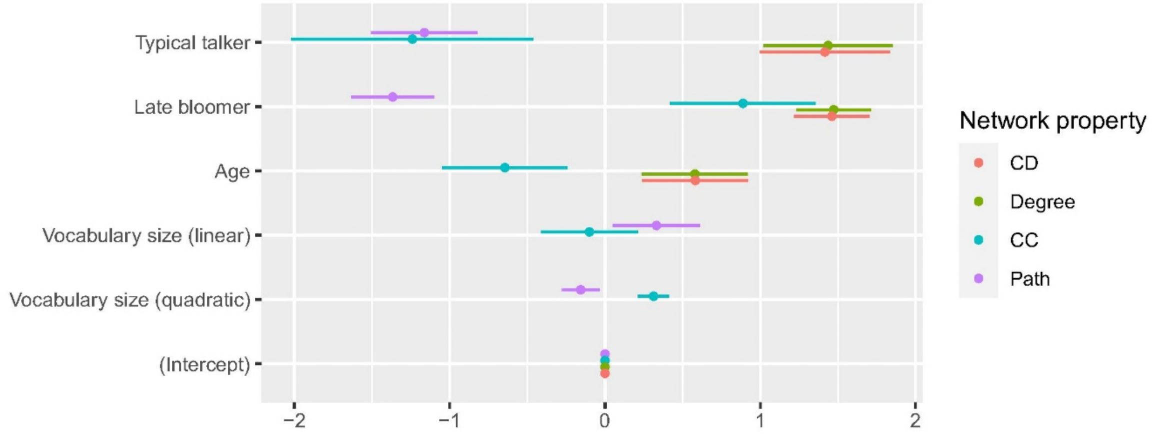
Contextual Diversity and other Network Properties in the Speech Directed to Late Talkers, Typical Talkers and Late Bloomers.



Note. LT= late talkers, TT= typical talkers, LB= late bloomers. Each dot represents a set of documents randomly sampled from either the LT corpus (blue), the TT corpus (orange) or the LB corpus (gray). The contextual diversity, age, and vocabulary size from each set of documents were averaged. The long regression line was fitted on all the data, and the short regression lines were fitted on each talker group. Shadows around the lines represent 95% confidence intervals.

Figure 5

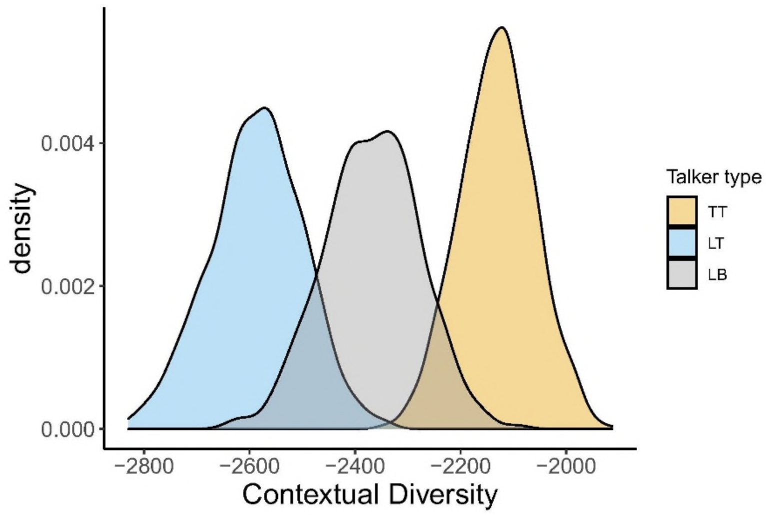
Standardised Regression Coefficients of Best Models Predicting Contextual Diversity and other Network Properties



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.

Figure 6

Contextual Diversity of in the Speech Directed to Late Talkers, Typical Talkers and Late Bloomers after Controlling for Frequency and Semantic Density



Note. Tukey's Ladder of Powers was conducted to transform values to comply with regression's assumptions.

Are the network properties of child-directed speech correlated with the network properties of children's vocabularies?

We carried out two sets of analysis to find out whether there is a correlation between the network properties of the children's vocabularies and the network properties of child-directed speech directed to the different talker-types, either with the same age (analysis 1) or the same vocabulary size (analysis 2). Spearman correlations displayed a significant relation between the network properties of the child-directed speech and the children's vocabularies. These relationships showed the same directionality for both our analyses using age- and vocabulary-size-matched language environments. We found a negative correlation for contextual diversity (vocabulary size: $p < .001$, $r = -.61$; age: $p < .001$, $r = -.34$), and positive correlations for mean degree (vocabulary size: $p < .001$, $r = .76$; age: $p < .001$, $r = .23$), clustering coefficient (vocabulary size: $p < .001$, $r = .51$; age: $p < .001$, $r = .52$) and average path length (vocabulary size: $p < .001$, $r = .22$; age: $p < .001$, $r = .47$). The negative correlation between contextual diversity in the environment and the vocabularies' contextual diversity confirm two pieces of evidence in the literature: that contextual diversity increases in parental speech as children age (Hills, 2013) and that children learn high contextual diversity words first (Hills et al., 2010). With regards to network measures, our correlation results suggest that the network properties of a child's lexicon correspond to the speech that they hear as they age and their vocabulary develops.

Do properties of the language environment predict vocabulary growth?

We also sought to find out whether the talker-type network metrics for the language environment could predict the vocabulary growth that children in our sample experienced during the study. With exception of clustering coefficient (p>.05), all the other networks properties predicted vocabulary growth. Vocabulary growth increases when contextual diversity and mean degree increases and average path length decreases in the language input (Diversity: b= 15.7 , F(1, 61)= 25.92, $p< .001$, Adj.R^2= .29; Degree: b= 15.6 ,F(1, 61)= 26.31, $p< .001$, Adj.R^2= .30; Path: b= -2353.5, F(1, 61)= 28.3, $p< .001$, Adj.R^2= .3;)

What are the contributions of strategy and environment on lexical development?

Figure 7 shows the growth of average degree (top), clustering coefficient (middle), and average path length (bottom) of the receptive (left-hand side) and expressive (right-hand side) lexical networks of LTs, TTs, and LBs. The same adjacency matrix (produced from the whole corpus) was used to build the children's networks. We refer to this as the 'same environment' in the statistical analysis. 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 5 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 their predictive power over a simpler version of the model, which does not have talker type as a predictor, indicated in Table 5 with an asterisk. When using the aggregate language environment to construct each child's network (i.e., same-environment analysis, the top half of Table 5), 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. On the other hand, network properties of children's expressive vocabularies were found to be similar across talker types. No differences across talker types were found for mean degree. These results suggest that, based on the aggregate learning environment, the different talker types learn structurally different words earlier for their receptive vocabularies but learn structurally more similar words for their expressive vocabularies.

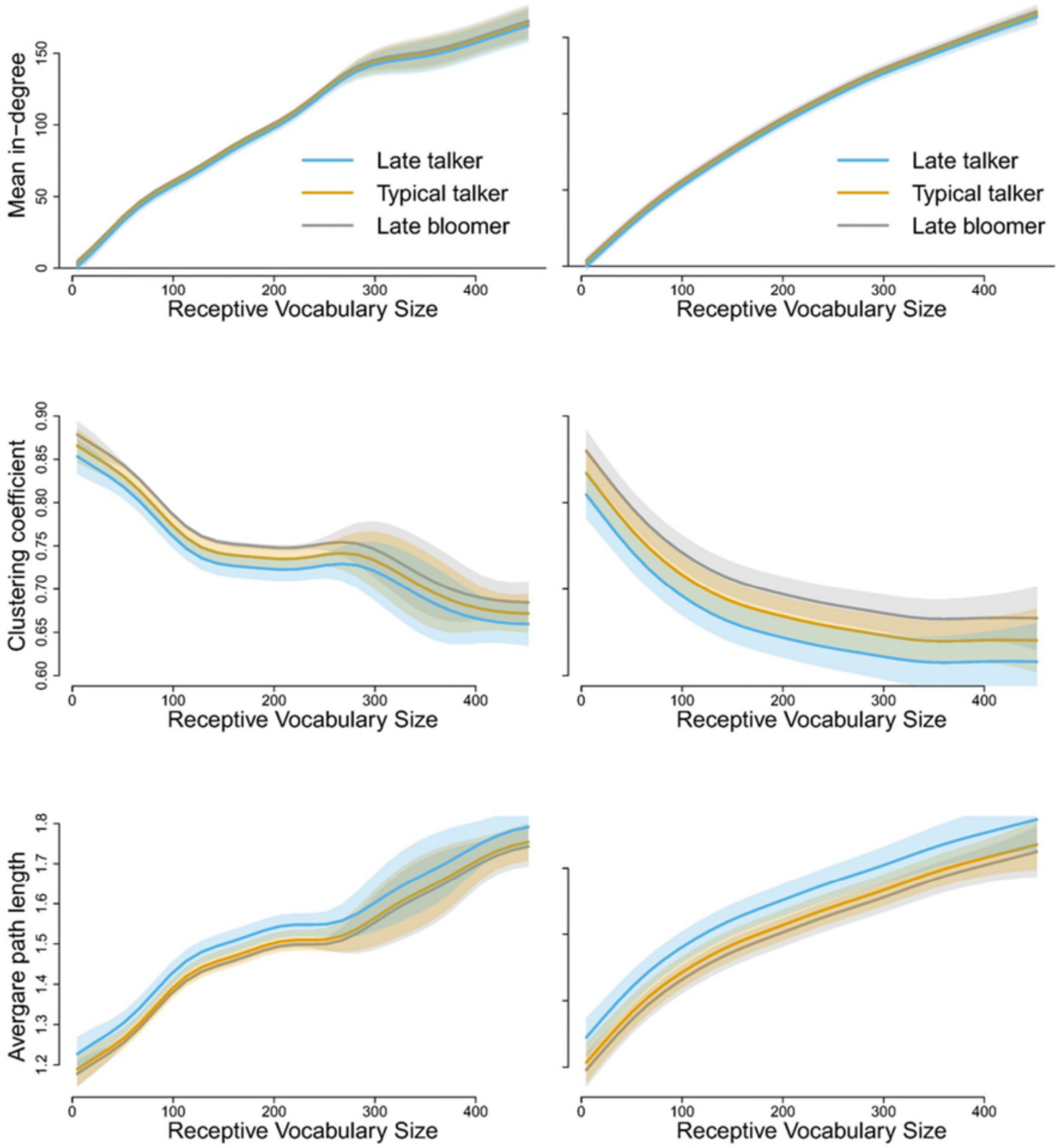
Figure 8 shows the network properties of children's vocabularies when the networks were built using their corresponding talker-type-related adjacency matrix. We refer to this as 'different-environment' in the statistical analysis. Results from our analysis are shown in the bottom half of Table 5. In this analysis, LBs showed higher mean degree and higher clustering coefficient than LTs in their receptive vocabularies; however, neither of these two models improved the predictive power of simpler models after including talker type. Average path length was identical across talker types. However, 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. These results suggest that the talker-types are using similar strategies to learning words with similar network properties in their receptive vocabularies (i.e., learning similar network structures), but because they are in different environments, these words appear different in the aggregate ('same environment') analysis. For their expressive vocabularies, the results based on their talker-type environment ('different environments') indicate the different talker-types are using different strategies (i.e., learning words with different networks properties).

The differences in the X and Y axis between Figure 7 and Figure 8 are dependent on differences between the corpora associated with each type of analysis. That is, "same-environment" analysis shows larger scales because the whole corpus was used to generate the adjacency matrix, and more words in the children's vocabularies could be included in the analysis. In contrast, in "different-environment" analysis, we carried out a bootstrapping procedure which reduced the amount of data used to generate the adjacency matrices. This was because we only considered the words that the three corpora had in common.

In sum, the evidence indicates that LTs, TTs, and LBs learn a different set of words, with contributions both from different learning strategies for the expressive vocabulary—inferred from the 'different environment' analysis—and different environments for the receptive vocabulary—inferred from the comparison between same and different environments.

Figure 7

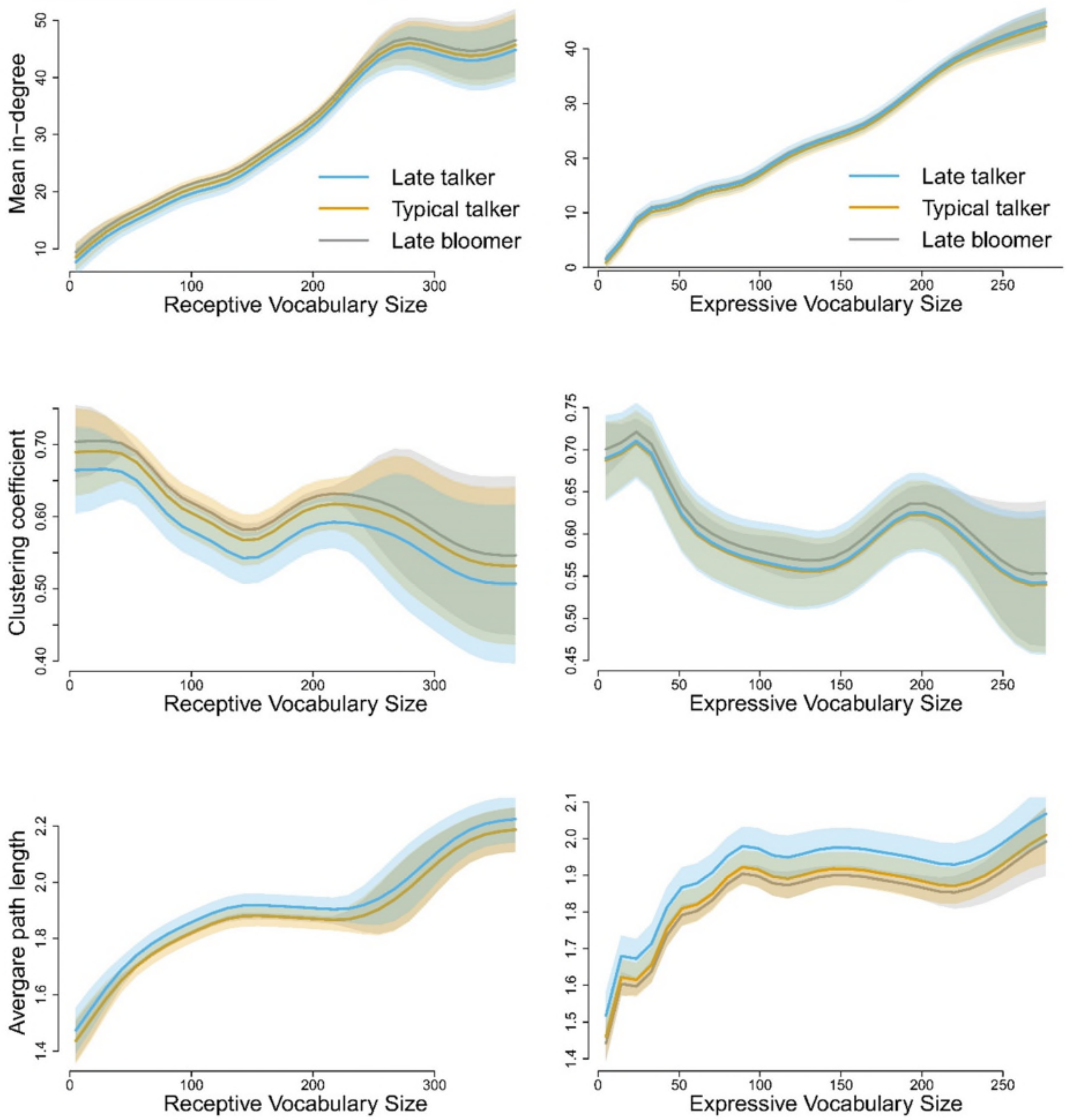
Network Properties of Growing Vocabularies of Typical Talkers, Late Talkers, and Late Bloomers Built from the Same Adjacency Matrix.



Note. The same adjacency matrix was generated from the whole corpus. Smoothed plots are based on the GAM models predictions. Shadows around the curves represent 95% confidence intervals.

Figure 8

Network Properties of Growing Vocabularies of Typical Talkers, Late Talkers, and Late Bloomers Built from Different Adjacency Matrices.



Note. Networks 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 5

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

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

Note. * indicates a significant main effect of the 'type of talker' variable for the model. Adding an interaction between talker type and vocabulary size did not significantly improve the prediction power of any model. All models had vocabulary size as a smoothing term in the GAM. CD= contextual diversity. Degree= average degree. CC= clustering coefficient. Path= average path length.

Discussion

The present study reveals a relation between the semantic structure of children's language environment and the semantic structure of their vocabulary and developmental trajectory. We show that the semantic richness and network properties of child-directed speech change over the first stages of children's lexical development and that children's network vocabularies reflect these changes. Further, the structural quality of the child-directed speech experienced by late talkers (LTs), late bloomers (LBs), and typical talkers (TTs) is different, and part of these differences correspond with the vocabulary structures of LTs, TTs, and LBs. In what follows, we discuss these results in more detail.

The Language Environment and Children's Vocabularies

The current study is the first to demonstrate that the network structure of parental speech directed to children between one and three years of age change over time. In both environment and children's vocabularies, the mean degree and average path length increase with time whilst clustering coefficient decreases. The fact that network properties of the environment and the children's vocabularies simultaneously change over time, supports a tight linkage between the environment and the learning mechanism. This is consistent with the growth principle of 'preferential acquisition' (Hills et al., 2009; Amatuni & Bergelson, 2019), but our analysis is not sufficient to determine causality.

Our results show a rising pattern of exposure to contextual diversity in child-directed speech as children age. This suggests that the vocabulary size of the child matters to the parents, who might tend to adapt the types of words they choose to expose to their children. Parents could introduce words with relatively low contextual diversity earlier (i.e., words with low mean degree, relative to words they produce later), and then when they notice a growth in their child's language abilities, their choice of words changes to include more words with higher contextual diversity. This parental adaptation to children's language skills is consistent with previous studies (e.g., Dykstra et al., 2012;

Hani, Gonzalez-Barrero, & Nadig, 2013; Paul & Elwood, 1991). Our results also confirmed the same increase of contextual diversity in parental speech captured in Hills (2013).

The early exposure of low-contextual-diversity words may suggest that a certain level of consistency (i.e., low contextual diversity) at early developmental stages facilitates word acquisition. We speculate that if we had collected audio data when participants were infants, we would have found even lower levels of contextual diversity in parental speech. Our results suggest that different levels of diversity might play different roles in word learning and that these roles are more appropriate at different stages of children's early life. We hinted in the introduction that consistency might assist in the process of learning the labels for referents through repetition, whereas contextual diversity might assist more in the process of learning the word meaning or semantic enrichment. If this hypothesis is true, it would also suggest that children generally learn the label first and then the semantic knowledge of the referent/word. Future experimental research is needed to investigate these speculative explanations.

The environment of late talkers, typical talkers, and late bloomers

The language environments experienced by LTs, TTs, and LBs present different degrees of semantic richness. Nonetheless, contextual diversity followed the same pattern of growth across the three talker-types: the larger the vocabulary size, the more contextual diversity the child experienced in the language input. Once corpus frequency and semantic density are controlled for, we found the children who presented the lowest vocabulary growth during the study, i.e., the LTs were exposed to the poorest language input, semantically speaking. This contrasts with the richer language input experienced by those children who showed the highest vocabulary growth during the study, i.e., the TTs and LBs. Interestingly, TTs received the richest language input of the three groups. These environmental differences align with the group’s language abilities: the richer the language environment, the better the language outcome. This is also supported by our correlational results which revealed an association between vocabulary growth and the amount of semantic richness in the language heard by the children. The outcome of the current study is consistent with the idea that contextual diversity has an important role as a facilitator for word learning. Although the present study cannot confirm that semantically poor parental speech causes early language delay, our strong evidence for an association is a first step into investigating this potential causal link.

The network properties of the parental speech also differed across the talker-types. The best language outcomes, i.e., TTs, are associated with a language environment whose network structure has high connectivity (degree), low semantic clustering, and low distance between words. The network properties in the parental speech directed to LBs resembles that of TTs, with the exception of the level of semantic clustering. Given that TT is the most successful group in terms of lexical development, it might be the case that TT families generate linguistic situations that increase the opportunities for word learning, and these linguistic situations are characterized as generating low clustering in the structure of their speech and having a short semantic distance between words. With respect to semantic clustering (i.e., clustering coefficient), TT families might delay producing speech that could make their children semantically associate groups of words among themselves, such as *baby-bottle-milk*. The late acquisition of these clusters of words (i.e., understood words) might facilitate word production in early stages of language development. Interestingly, the parental speech of LB families showed the highest clustering coefficient, even higher than that of LT families, which questions the hypothetical role of semantic clustering in child-directed speech on word production. The mismatch between language outcome and clustering coefficient could explain why clustering did not predict vocabulary growth. With regards to semantic distance between words (i.e., average path length), the families of TTs and LBs also produced speech that might encourage children to create stronger semantic associations between words compared to LT families. For example, a TT could find *baby* and *break* to be semantically related because, from what she learned from the speech she hears, these two words are linked through the word *bottle* (*baby*→*bottle*→*break*), whereas a LT finds a weaker semantic association between these two words because in her language environment they are connected by two words (*baby*→*bottle*→glass→*break*). Although no research has been conducted to prove how these network structures in parental speech might promote (or slow down) vocabulary growth, the current results warrant further investigation.

The quantity versus quality debate has been a recurrent topic in the language development field since the publication of the Hart's and Risley's study in 1995. The present study has isolated the effect of quality from quantity in various ways (i.e., sampling the same number of words from each talker-type corpus, normalizing contextual diversity by dividing the word's contextual diversity by its frequency in the sample, and dividing this value by the total audio duration). The results showed that children with different language outcomes experienced a language environment with different degrees of quality (measured by contextual diversity), where typical talkers received the richest input, followed by late bloomers, and finally late talkers received the poorest input. With respect to quantity, TTs experienced a higher quantity of language than LBs and LTs (measured by semantic density), with LTs and LBs experiencing a similar amount. In light of these results, an interesting pattern is observed: TT environments are rich in quantity and quality; LB environments are rich in quantity but poor in quality; and LT environments are poor in both quantity and quality. What can be infered here is that: 1) both quantity and quality may promote language development; and 2) LBs might be initially delayed by the low quantity in the language input. Further research will be needed to confirm the causality of these interesting inferences.

Strategy and Environment in Vocabulary Development

There are two main ways in which children might differ in the network properties of their same-size vocabularies: by learning different types of words as a result of different learning strategies or by learning different types of words as a consequence of experiencing different language environments. For receptive vocabularies, we found that LTs understand different types of words that make their receptive networks display lower semantic clustering and higher path length compared to TTs and LBs. However, when we considered the semantic differences from their respective talker-type environments (i.e., 'different-environment analysis'), the differences in the vocabularies between the talker-types disappeared. These results suggest that, for receptive vocabularies, the different talker-types are using similar strategies but developing in different language environments. This is, LTs, TTs, and LBs appear to understand different words, but these words have similar degrees of connectivity in their respective environments. In this way, the semantic structure of the environment contributes to the lexical development of the child’s receptive vocabulary.

For expressive vocabularies, we found that LTs produce structurally similar words to TTs and LBs when using the aggregate shared-edge rules. The absence of differences between LTs and TTs is inconsistent with previous findings (Beckage et al., 2011; Hills et al. 2010; Jiménez & Hills, 2017), but here we are using a different data set and different analyses (e.g., including longitudinal data and not using a random comparison group). Nonetheless, the analyses based on idiosyncratic learning environments find a difference between talker types, i.e., higher average path length for LT children. In contrast to receptive vocabularies, the words that each talker group produces have different degrees of connectivity in their respective environments. Although this might suggest that the LT children are using a different learning strategy in their respective language environments, the trajectory/curve of the average path length (and other network properties) is similar across talkers. What could be suggested instead is that LTs use a similar strategy but a weaker one. This is different to the interpretation made by Beckage et al. (2011) which suggested that LT children may choose to learn words that are less similar to words they already know. In network science, higher average path length translates into a less efficient distribution of information throughout the system at a global level (Barabási 2016), which may also apply during lexical development. In other words, the higher distance between words in LTs vocabularies might have consequences when processing language, e.g., it might take longer for them to navigate from one concept to another.

In sum, our findings suggest different results between expressive and receptive vocabularies: LTs, TTs, and LBs understand a different set of words because they experienced a different language input, and produce different sets of words because they use word learning strategies at different degrees. Future research is needed to disentangle these results, perhaps focusing on how words are promoted from receptive vocabularies to expressive vocabularies whilst taking into account how words are connected in the receptive vocabulary.

Limitations and 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 autism spectrum disorder, 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 developmental language disorder in the future. Also, the vocabulary norms utilized to identify the children's percentiles come 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. Also, 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. Furthermore, as we indicated in the Methods section, there were not enough transcriptions per family to conduct a more fine-grained network analysis of child-directed speech at the level of individual children, 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. We also found that low levels of clustering coefficient and path length are associated with the best language outcomes. Future research could characterize the type of linguistic situations that lead to a language structure high in contextual diversity but relatively low in clustering coefficient and path length and how this linguistic structure promotes word acquisition. Finally, our data does not involve a controlled experimental manipulation and therefore does not allow us to make strong causal inferences. Children may adapt to their parents' speech, parents may adapt their speech to their child's developmental trajectory, or, more likely, both. Future research will be needed to tease apart the causal pathways. The present work nonetheless demonstrates a clear relationship between language development and environment.

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

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

Amatuni, A., & Bergelson, E. (2017). Semantic networks generated from early linguistic input. bioRxiv, 157701.

Barabási, A.L. (2016) *Network Science.* Cambridge University Press.

Beckage, N. M., & Colunga, E. (2019). Network growth modeling to capture individual lexical learning. Complexity, 2019.

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

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

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.

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

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

Ellis Weismer, S. (2007). 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

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

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.

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

Green, P., & MacLeod, C. J. (2016). SIMR: an R package for power analysis of generalized linear mixed models by simulation. Methods in Ecology and Evolution, 7(4), 493-498.

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

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

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. (2009). 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

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

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

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., 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

Johns, B. T., Dye, M., & Jones, M. N. (2016). The influence of contextual diversity on word learning. Psychonomic bulletin & review, 23(4), 1214-1220; Johns, B. T. (2021). Disentangling contextual diversity: Communicative need as a lexical organizer. Psychological Review, 128(3), 525.

Markman, E. M. (1991). 3. The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints. *Perspectives on language and thought: Interrelations in development*, 72.

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

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

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

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

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).

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

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

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

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

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

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

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., & 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

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

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

Schwab, J. F., & Lew-Williams, C. (2016). Language learning, socioeconomic status, and child-directed speech. Wiley Interdisciplinary Reviews: Cognitive Science, 7(4), 264-275.

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

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

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

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., 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

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

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

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