**Late Bloomer or Language Disorder? Differences in Toddler Vocabulary Composition Associated with Long-term Language Outcomes**

Lynn K. Perry1

Sarah C. Kucker2

Jessica S. Horst3

Larissa K. Samuelson4

1University of Miami

2Oklahoma State University

3University of Sussex

4University of East Anglia

**Author Note**

Lynn K. Perry https://orcid.org/0000-0001-6976-3741

Sarah C. Kucker https://orcid.org/0000-0003-2210-3599

Jessica S. Horst https://orcid.org/0000-0003-3200-5239

Larissa K. Samuelson https://orcid.org/0000-0002-9141-3286

**Conflict of Interest:** The authors have no conflicts of interest.

**Data Availability:** The data that support the findings of this study are available on the Open Science Framework: https://osf.io/yjstq/ and on Wordbank: wordbank.stanford.edu

**Acknowledgements:** This research was funded by a grant from NICHD #R01HD045713 awarded to LKS. We thank the members of the Iowa Language and Category Development Lab, Sussex Word and Object Reasoning Development (WORD) Lab, Oshkosh Development Lab, and Miami Object and Word Learning (OWL) Lab for help in data collection and entry. We also thank Donna Thal and Bruce Tomblin for making diagnosis data available and Michael Frank for sharing R code for processing Wordbank data.

**Author Contributions:** LKP and JSH conceived the original study idea and curated data. All authors were involved in data collection. LKP designed analyses and processed and analyzed the data. LKP and SCK wrote the first draft of the manuscript. All authors edited the manuscript.

**Corresponding Author:** Correspondence concerning this article should be addressed to Lynn K. Perry, Department of Psychology, University of Miami, 5665 Ponce de Leon Blvd., Coral Gables, FL, 33146. Email: lkperry@miami.edu

**Research Highlights**

* Toddler vocabulary composition, including the proportion of names for categories organized by shape, like spoon, was used to retrospectively compare outcomes of late talking children
* Persisting Late Talkers said a smaller proportion of shape-based nouns during toddlerhood relative to Late Bloomers (late talkers who later caught up to have typically-sized vocabularies)
* Children with later DLD diagnoses said a smaller proportion of shape-based nouns during toddlerhood relative to children without a DLD diagnosis
* The data illustrate the cascading effects of vocabulary composition on subsequent language development and suggest vocabulary composition may be one important marker of persisting delays

**Abstract**

Children with delays in expressive language (late talkers) have heterogeneous developmental trajectories. Some are late bloomers who eventually “catch-up”, but others have persisting delay or are later diagnosed with developmental language disorder (DLD). Early in development it is unclear which children will belong to which group. We compare the toddler vocabulary composition of late talkers with different long-term outcomes. The literature suggests most children with typical development (TD) have vocabularies dominated by names for categories organized by similarity in shape (e.g., cup), which supports a bias to attend to shape when generalizing names of novel nouns—a bias associated with accelerated vocabulary development. Previous work has shown that as a group, late talkers tend to say fewer names for categories organized by shape and are less likely to show a “shape bias” than TD children. Here, in a retrospective analysis of 850 children, we compared the vocabulary composition of groups of toddlers who were late bloomers or persisting late talkers. At Time 1 (13-27 months), the persisting late talkers said a smaller proportion of shape-based nouns than both TD children and late bloomers who “caught up” to typically sized vocabularies months later (18-38-months). Additionally, children who received a DLD diagnosis between 4 and 7 years said a significantly smaller proportion of shape-based nouns in year two than TD children and children with other diagnoses (e.g., dyslexia). These findings bring new insight into sources of heterogeneity amongst late talkers and offer a new metric for assessing risk.

*Keywords*: late talker; developmental language disorder; vocabulary; shape bias

**Late Bloomer or Language Disorder? Differences in Toddler Vocabulary Composition Associated with Long-term Language Outcomes**

 Children’s language skills and vocabulary knowledge rapidly increase during the second year of life. By the time they are 24 months old, the majority of children have an expressive vocabulary of between 150 and 450 words (Fenson et al., 1994). However, there are vast individual differences in children’s vocabulary development such that many children do not make such rapid gains and have much smaller expressive and/or receptive vocabularies than average. These children are commonly referred to as “late talkers” (Macroy-Higgins & Montemarano, 2016; Rescorla, 2011). Late talkers vary from each other in their vocabulary size, phonological skills, pragmatic and social abilities, demographic backgrounds, and more (see Desmarais et al., 2008). Moreover, late talkers also vary greatly in their long-term outcomes. Around half of these children are “late bloomers,” eventually catching up to their peers in the following years and having vocabularies in the normal range (Rescorla, 2011; Singleton, 2018). Nevertheless, a sizable number of late talkers continue to have persisting language delays and are later diagnosed with Developmental Language Disorder (DLD; Bishop, 2017; Rescorla, 2011; McGregor et al., 2020). As DLD has significant cascading impacts on life outcomes, knowing which late talkers will bloom and which will continue to have delays is critical for targeting early intervention (Singleton, 2018). Unfortunately, the field still lacks an understanding of the early predictors of persisting delay or DLD (Rescorla, 2011; Samuelson, 2021). For example, although vocabulary size is one metric for assessing delay, early vocabulary size alone does not clearly predict later delay (Duff, Nation, et al., 2015; Duff, Reen, et al., 2015). Here we consider whether differences in vocabulary composition, that is, the specific words children say, are associated with long-term outcomes.

**Late Talking Children**

Late talkers are children who show significant delays in their expressive vocabulary development relative to age-based norms (Bishop, 2017; Dollaghan, 2013; Ellis Weismer, 2007; Rescorla, 2011). A large body of work has been dedicated to characterizing this group with respect to who is most likely to have persisting delays and the nature of those delays. For example, boys are more likely to be late talkers than girls (Armstrong et al., 2017), preterm infants are more likely than full term infants (Rudolph, 2017; Zubrick et al., 2007), younger siblings are more likely than older siblings (Rudolph, 2017), and children from lower socioeconomic households are more likely to be late talkers than those from higher SES backgrounds (Armstrong et al., 2017; Rescorla, 2011; Rudolph, 2017). In addition to delays in vocabulary, late talkers are likely to have delays in other aspects of language and learning. Children who are late talkers have slower speech processing speed (Ellis et al., 2015; Fernald & Marchman, 2012), difficulty with referent selection (Kucker & Seidler, 2022) and retaining new words (Ellis Weismer et al., 2013), delays in lexical growth (Rescorla, 2000), poor oral narrative skills (Manhardt & Rescorla, 2002), and poor social abilities (Longobardi et al., 2016). Late talkers are also less likely to recognize familiar objects’ shapes (Jones & Smith, 2005) or selectively attend to visual information like shape as they learn new words (Jones, 2003). However, despite the wealth of studies characterizing the late talking population in toddlerhood, few studies have examined vocabulary-level factors associated with long term outcomes like persisting delay and DLD diagnoses—a primary goal of this paper.

***Predicting Outcomes Among Late Talkers***

Some late talking toddlers continue to have delays throughout childhood and up to 40% go on to be diagnosed with Developmental Language Disorder[[1]](#footnote-2) (DLD; Rescorla, 2011; Bishop, 2017). DLD is characterized by poor language skills in the absence of other major causes. It is a high prevalence disorder, affecting an estimated 7% of the population (Tomblin et al., 1997; and see McGregor, 2020). Because language is deeply integrated into everyday life, DLD has cascading consequences for children’s academic achievement, high school graduation, incarceration rates, and more (Botting & Conti-Ramsden, 2008; Johnson et al., 2010; Le et al., 2020). Early identification of children at risk of persisting delay, including DLD, can help us better design and implement early interventions tailored toward individual children (Samuelson, 2021).

However, expressive vocabulary size in toddlerhood and late talker status alone are not consistently predictive of later outcomes (Duff, Nation, et al., 2015). Instead, differences in potential mechanisms by which toddlers learn language tend to be associated with their subsequent vocabulary development. For instance, 2-year-old late talking children’s sensitivity to phonological properties and novel word mapping abilities positively correlate with language delay at 3.5 years (Ellis Weismer, 2007). Similarly, 18-month-old children who are more efficient word processors (e.g., look to the referent of a spoken word faster and more accurately) show larger gains in vocabulary over the next year than less efficient word processors (Fernald & Marchman, 2012).

Notably, however, the factors predicting short-term delay over the course of a year (e.g., phonological sensitivity, word processing speed) do not necessarily predict longer-term persisting delay over many years or later DLD diagnosis (Rescorla, 2011). Likewise, delayed achievement of language milestones does not necessarily predict long-term delay. For example, although a delay in children’s ability to combine words (i.e., produce multi-word utterances) is one reliable predictor of later DLD diagnosis at school age, a delay in children’s first word production is not (Rudolph & Leonard, 2016). Additionally, some work suggests being in the lower end of expressive vocabulary at 24-months predicts long-term delay and/or diagnosis (Rescorla, 2009). Yet other work suggests that some children who eventually receive a diagnosis of DLD did not have any risk factors and did not demonstrate early markers of delay (Rescorla, 2000; Rescorla et al., 1997). The variability in late talkers’ long-term diagnostic and linguistic outcomes makes identifying robust early predictors of later delay difficult. One promising direction for identifying predictors of persisting delay in late talkers is to look at the processes that support language growth in typically developing children.

**The Importance of Vocabulary Composition in Development**

A wide range of research shows that infants and toddlers are adept at finding and using regularities to learn words, including statistical differences in transition probabilities (Graf Estes et al., 2007), patterns of word-object co-occurrence (Smith & Yu, 2008), and the relation between the syntactic context of a naming event (e.g. “this is a…”) and attention to specific object features (Landau et al., 1992). It has been suggested that picking up on this last regularity—between syntax and the perceptual properties of objects—means that regularities found in the vocabulary itself may be one important support for early vocabulary development. However, there are differences in the regularities presented by different languages and even in the specific words that individual children learning the same language say, meaning this support may not be equal for all children.

On average, children learning English acquire an early noun vocabulary that is dominated by words naming solid objects in categories organized by similarity in shape (e.g., ball, cup; Samuelson & Smith, 1999; Colunga & Smith, 2005; Gershkoff‐Stowe & Smith, 2004; Samuelson & Smith, 2000). As these children’s vocabularies grow, they become biased to attend to shape as they learn new words (i.e., show a “shape bias”). In this way, regularities in the vocabulary teach these children *how* to learn more efficiently in the future and support rapid vocabulary growth (Smith et al., 2002). However, some children do not learn as many of these “shape-based” words, and instead learn more of other types of words, such as those naming categories organized by similarity in material (e.g., chalk, ice; Perry & Samuelson, 2011). These differences in vocabulary are associated with differences in children’s likelihood of showing a shape bias versus a material bias (Perry & Samuelson, 2011), recognition of familiar objects (Perry & Saffran, 2017), and memory for specific features of novel objects (Perry et al., 2016; Slone & Sandhofer, 2017). Thus, although overall vocabulary *size* is associated with language processing skills (Fernald et al., 2006), which has consequences for subsequent vocabulary growth (Fernald & Marchman, 2012), a growing body of work demonstrates that the *composition* of a child’s vocabulary can be illustrative of the processes that support vocabulary growth (Perry & Samuelson, 2011), and, that vocabulary composition may be a particularly critical factor in identifying children with language delays (Ellis Weismer et al., 2011; Jiménez et al., 2021).

***Differences in Word Learning Biases and Vocabulary Composition among Late Talkers***

A growing body of evidence suggests that children with language delays, including late talkers (Colunga & Sims, 2017; Jones, 2003; Perry & Kucker, 2019), older children with language impairment (Collisson et al., 2014), and children with other delays related to hearing loss (Perry et al., 2021) or autism spectrum disorder (Potrzeba et al., 2015; Tek et al., 2008) do not show a shape bias at the same age as their typically developing peers. In fact, late talkers have particular difficulty selectively attending to the shape of objects. For example, they struggle with recognizing highly familiar objects from Styrofoam shape caricatures relative to both age-matched and younger vocabulary-matched peers with typically-sized vocabularies (Jones & Smith, 2005), suggesting potential differences in the way they visually process objects (Borgström et al., 2015, 2019), which is likely to have cascading effects on later vocabulary (Cf. Smith, 2009).

Critically, there appear to be differences in the underlying composition of late talkers’ vocabularies that are related to these differences in word learning and object recognition. In particular, there appear to be differences in the *types* of words said by late talkers compared to typically developing children. For example, compared to typically developing children, late talkers are more likely to know more oddball words that are not semantically related to each other, e.g., they are more likely to know “pool” and “scissor” than “hair” and “comb.” (Beckage et al., 2011; but see Jiménez & Hills, 2017). Such differences in the interrelatedness of the words children know are associated with subsequent differences in the rate of vocabulary development (Beckage et al., 2011) and likelihood of later DLD diagnosis (Borovsky et al., 2021). Late talkers who have persisting delay also know significantly fewer nouns than typically developing children (Weber & Colunga, 2021; and see MacRoy-Higgins et al., 2016; Jiménez et al., 2021). However, late talking children are more similar to typically developing children in the amount and types of verbs they know (Horvath et al., 2019), suggesting that differences in their vocabulary specifically with respect to nouns may be most critical.

As a group, however, late talkers vary greatly in the number of nouns they say that name categories organized by shape, while typically developing children are more similar to each other regardless of vocabulary size (Colunga & Sims, 2017; and see Perry & Kucker, 2019). Late talkers, and neural networks trained on the composition of those late talkers’ vocabularies, develop more heterogeneous word learning biases than early talkers (Colunga & Sims, 2017). We know that the homogeneity found in the vocabulary itself—typically in the form of many object nouns naming categories organized by similarity in shape—supports subsequent vocabulary development (e.g., Samuelson, 2002). Consequently, a heterogeneous vocabulary characterized by words that are neither semantically related nor dominated by categories organized by shape would not give children the regularity and support they need to learn how to learn new words, potentially leading some late talkers to have persistent language delays. A deeper understanding of this heterogeneity in the composition of late talkers’ toddler vocabularies may shed light on why some late talkers continue to have persisting language delays while other late talkers are able to catch up to their peers by the early school years. Here we focus specifically on the composition of late talkers’ toddler noun vocabularies.

**Current Study**

The primary goal of the current study is to assess which aspects of vocabulary composition differ systematically across late talkers who continue to have delays and those who catch up. In Study 1, we sorted 850 children into three groups based on their late talker status at two time points (Time 1 *M=*16.30 months and Time 2 *M=*27.48 months): (A) “Persisting Late Talkers” were children who were late talkers at Time 1 and who continued to be late talkers several months later at Time 2; (B) “Late Bloomers” were children who were late talkers at Time 1 but who had a typically-sized expressive vocabulary at Time 2; and (C) “Typical Talkers” were children who had typically-sized vocabularies at Time 1. We then used a retrospective approach to compare the groups with respect to their Time 1 vocabularies. Following previous work linking vocabulary and attention to shape, we used measures of expressive, rather than receptive, vocabulary (i.e., the MBCDI; see e.g., Colunga & Sims, 2017; Samuelson & Smith, 1999; Smith et al., 2002). We examined differences in the types of nouns in children’s expressive vocabularies, focusing in particular on the proportion of shape-based nouns[[2]](#footnote-3). We hypothesized that the Persisting Late Talkers would have a significantly smaller proportion of shape-based nouns in their Time 1 vocabularies than the Late Bloomers or Typical Talkers. Finally, we explored whether shape-based nouns not only differed across groups, but whether they were predictive of later outcomes in a complementary analysis in which we predicted outcome group from shape vocabulary.

In Study 2, a follow-up exploratory study, we examined the vocabularies of a subset of the children from Study 1 for whom long-term diagnosis data at 4-7 years was available. We sorted children into groups based on diagnosis, “DLD,” “Other Diagnosis” (dyslexia, speech disorder, or learning disability not in conjunction with DLD), or “No Diagnosis.” We used a retrospective approach to compare these groups with respect to their Time 1 vocabularies. We examined differences in the types of nouns in children’s expressive vocabularies at Time 1, particularly the proportion of shape-based nouns. We hypothesized that children who go on to receive a DLD diagnosis would have had a smaller proportion of shape-based nouns in their Time 1 vocabularies than those with Other Diagnoses or No Diagnosis. Finally, we again explored whether shape-based nouns not only differed across diagnosis groups, but whether they were predictive of later diagnosis in a complementary analysis predicting diagnosis group from shape vocabulary.

**Study 1**

**Methods**

***Participants***

Participants included 850 children (Time 1: *M=*16.30 months, range: 13-27 months, Time 2: *M=*27.48 months, range=18-38; 394 girls) whose parents had completed two *MacArthur-Bates Communicative Developmental Inventory: Words and Sentences* (MBCDI) forms (Fenson et al., 1994) at least three months apart (*M*gap length *=* 10 months). Of these, 197 (Time 1: *M =* 17.29 months, *SD* = 3.30 months; Time 2: *M =* 25.75, *SD* = 3.59; 89 girls) children had participated at the University of Iowa (*N =* 109) or the University of Wisconsin Oshkosh (*N =* 88) in two separate lab studies using the MBCDI to measure expressive vocabulary knowledge. The remaining 653 children (Time 1: *M =* 16.00 months, *SD* = .06 months; Time 2: *M =* 28.00, *SD* = .04; 305 girls) had participated in Thal and colleagues’ (2013) longitudinal study and their MBCDI data were shared on WordBank (Frank et al., 2017).

***Group Characterization***

Using MBCDI normative data, we categorized children based on their total expressive vocabulary size at Time 1 as Late Talkers (< 25th percentile for age) or Typical Talkers (> 25th percentile for age). We then categorized the late talkers into two further groups: children who were still below the 25th percentile at Time 2 (Persisting Late Talkers) and those who were above the 25th percentile at Time 2 (Late Bloomers). We used the 25th percentile to have a broad sample of children at the lower end of our range, similar to other recent studies (e.g., Colunga & Sims, 2017; Perry & Kucker, 2019; Kucker & Seidler, 2022). Late Bloomers and Persisting Late Talkers did not differ in their vocabulary size at Time 1, *p = .*759, and both had significantly smaller vocabularies than Typical Talkers, *ps <* .00001. As can be seen in Table 1, Persisting Late Talkers tended to be slightly older than Late Bloomers and Typical Talkers at Time 1. We therefore include age at Time 1 as a covariate in our analyses.

**Table 1. Group characteristics**. Late talker status was defined as having a vocabulary size below the 25th percentile for one’s age. Ages and vocabulary sizes are listed as *M* (SD).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Group**  | **Sample****Size** | **Age at Time 1** | **Vocabulary size at Time 1** | **Age at Time 2** | **Vocabulary size at Time 2** |
| Typical Talkers | 557(278 girls) | 16.22 months(1.64 months) | 86.67 words(90.03 words) | 27.47 months(2.03 months) | 511.37 words(150.57 words) |
| Persisting Late Talkers | 106(36 girls) | 16.78 months(2.13 months) | 14.62 words(21.04 words) | 27.07 months(1.89 months) | 188.69 words(106.31 words) |
| Late Bloomers | 187(80 girls) | 16.27 months(1.43 months) | 17.48 words(22.12 words) | 27.75 months(1.80 months) | 496.01 words(113.79 words) |

***Vocabulary Size and Composition***

Because our primary measure of interest relates to the composition of children’s object noun vocabulary, we compare the number of object nouns in children’s vocabularies. Object nouns are those from Sections 2-10 of the MBCDI (animals, vehicles, toys, food and drinks, clothing, body parts, small household objects, furniture and rooms, and outdoor things). Noun vocabulary is a possible area of deficit in children with language delays (Horvath et al., 2019; Jiménez & Hills, 2017; MacRoy-Higgins et al., 2016) and critically, object nouns in particular have previously been shown to relate to children’s word learning skills (e.g., Samuelson, 2002; Samuelson & Smith, 1999).

We then considered *vocabulary* *composition* by examining object nouns (Sections 2-10) with respect to their dominant category features (shape, material, or multidimensional organization). Our primary measure of vocabulary composition involved calculating the proportion of nouns that adults judge to name categories organized by similarity in shape in children’s vocabularies at Time 1 (so called “shape-based” nouns, similar to Perry & Samuelson, 2011; Perry et al., 2016; Perry & Saffran, 2017; Slone & Sandhofer, 2017; Perry & Kucker, 2019) divided by the total number of object nouns in their vocabularies at Time 1. According to previous work using adult judgments (e.g., Samuelson & Smith, 1999), shape-based nouns include, for example, categories of solid objects organized by similarity in shape that use count syntax, such as “a,” “one,” or “two” (e.g., ball), categories of solid objects organized by similarity in shape that have no agreed upon syntax (e.g., green beans), and categories organized by similarity in shape that use count syntax, but have ambiguous solidity (e.g., sweater). According to adult judgments of the 312 object nouns on the MBCDI, 185 of them are shape-based (Samuelson & Smith, 1999). See Table 2 for descriptions and examples of these classifications.

We also assessed vocabulary composition with respect to the proportion of nouns that adults judge to name categories organized by material similarity, and nouns with a multidimensional category structure. The 50 “material-based” nouns[[3]](#footnote-4) on the MBCDI include nonsolid substances in categories organized by similarity in material that use mass syntax, such as “some” (e.g., applesauce), categories of nonsolid substances organized by similarity in material that have no agreed upon syntax (e.g., soda/pop), categories organized by material that use mass syntax, but have ambiguous solidity (e.g., butter), categories of solid objects organized by material similarity that use mass syntax (e.g., chalk), and categories organized by material that use count syntax (e.g., towel). The 82 nouns with multidimensional category structure are those where there is no single property that adults agree defines the category, including nouns such as bathroom, blanket, pizza, present, story, and wind. See Table 2.

**Table 2.** Descriptions and examples of each type of noun vocabulary classification (based on Samuelson & Smith, 1999 adult judgments; and see Perry & Samuelson, 2011).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Noun vocabulary measure** | **Sub-classification type** | **Description** | **Examples** | **Number of MCDI words**  |
| **Shape-based vocabulary** | Solid + Shape + Count | Solid objects in categories organized by similarity in shape that use count syntax | ball, dog, car  | 107 |
| Solid + Shape | Solid objects in categories organized by shape with no agreed upon syntax | corn, green beans, pants | 7 |
| Solid + Count | Solid objects that use count syntax | book, cracker, camera | 49 |
| Shape + Count | Categories organized by similarity in shape that use count syntax with ambiguous solidity | pancake, sweater, vagina | 24 |
| **Material-based vocabulary** | Nonsolid + Material + Mass | Nonsolid substances in categories organized by similarity in material | applesauce, jelly, water | 6 |
| Nonsolid + Material | Nonsolid substances in categories organized by material with no agreed upon syntax | Coke | 1 |
| Nonsolid + Mass | Nonsolid substances named by mass syntax | juice, sauce, soup | 5 |
| Material + Mass | Categories organized by similarity in material that use mass syntax | Play Doh, jello, yogurt  | 12 |
| Solid + Material + Count | Solid objects in categories organized by material named by count syntax | apple, lollipop, jeans | 10 |
| Solid + Material + Mass | Solid objects in categories organized by material named by mass syntax | cheese, ice | 2 |
| Solid + Material | Solid objects in categories organized by material with no agreed upon syntax | chalk, french fries, paper | 11 |
| Solid + Mass | Solid objects named by mass syntax | bread, meat | 2 |
| Material + Count | Categories organized by material named by count syntax | towel, snowman | 2 |
| **Multidimensional vocabulary** | All other object nouns | Categories not fitting into either the shape or material-based classifications | blanket, story, pizza, present, bathroom, wind | 74 |

**Results**

The goal of these retrospective analyses was to examine whether the groups (Typical Talkers, Late Bloomers, Persisting Late Talkers) differed in noun vocabulary size or vocabulary composition at Time 1; exploration of other lexical classes on the MBCDI are presented in Supplemental Information. First, we examined differences between the groups in noun vocabulary size. We then conducted chi-square analyses to examine whether there were overall group differences with respect to the number of words children said in each vocabulary classification (e.g., number of shape-based nouns versus material-based nouns versus multidimensional nouns). Next, to examine what aspect of vocabulary composition *specifically* might differ across groups, controlling for vocabulary size and other covariates like age and sex, we conducted multiple regression analyses in which we predict the proportion of children’s object noun vocabulary falling into a given classification (e.g., proportion of shape-based nouns) from group. Proportion variables were log transformed using log10 (x + .01) where x was the proportion. Regression analyses were conducted in R (R Core Team, 2014). Persisting Late Talkers were used as the reference group in these analyses. Bonferonni corrections were made to control for family-wise error related to multiple comparisons of proportions. Adjusted alpha levels are listed in each regression table. Finally, we performed logistic regression analyses predicting group from proportion of shape-based nouns (log transformed), controlling for age and sex.

***Differences in Object Noun Vocabulary Size***

We first examined differences in children’s object noun vocabulary size. Using a multiple regression model, we compared the total *number* of object nouns in children’s Time 1 vocabulary based on their Time 2 group (Typical Talkers, Persisting Late Talkers, Late Bloomers). The model also included age at Time 1 and sex as covariates due to previous findings that girls often have larger vocabularies and are less likely to be late talkers than boys (Dale & Fenson, 1996; Huttenlocher et al., 1991; Rudolph, 2017). Typical Talkers (*M =* 86.67 words, *SD* = 90.03) had significantly larger Time 1 object noun vocabularies than Persisting Late Talkers (*M =* 14.62 words, *SD* = 21.04), *B =* 48.48, se = 3.69, *t =* 13.14, *p <* .00001, who had similar noun vocabularies as Late Bloomers (*M* = 17.48 words, *SD* = 22.12). These findings are unsurprising given that the norms used to define groups are based on total vocabulary size which is highly correlated with object noun vocabulary size (*r* = .99 in the current sample; see also Perry & Saffran, 2017; Samuelson, 2002; Samuelson & Smith, 1999). There was only a marginal difference in object noun vocabulary size between the two late talker groups, *B =* 7.79, se = 4.20, *t =* 1.85, *p =* .064, confirming that something beyond the *number* of object nouns in the vocabulary of children within the bottom quartile is needed to explain some children’s movement out of that quartile. There was also a significant effect of sex on vocabulary size, *B =* 5.66, se = 2.37, *t =* 2.39, *p =* .017.

 As a result of producing fewer object nouns than children in the other groups, the vocabulary composition of Persisting Late Talkers included a higher proportion of words falling into the “other words” category (see Supplemental Information). These other words primarily consisted of sound effect words (i.e., onomatopoeia) and names of games and routines. We discuss this result and report other differences in vocabulary across lexical classes in the Supplemental Information.

***Differences in Vocabulary Composition***

To explore vocabulary composition, we first performed a chi-square analysis comparing the number of words said by each group in each classification and found significant group differences, *X2*(4) *=* 36.74*, p <* .00001.Given the group differences, we conducted separate multiple regression analyses to examine specific differences in vocabulary composition across groups.

**Shape-based Noun Vocabulary.** We explored vocabulary composition with respect to the proportion of *shape-based* nouns in children’s vocabularies. Using a multiple regression model, we compared the proportion of shape-based words in children’s Time 1 vocabulary based on their Time 2 group (Typical Talkers, Persisting Late Talkers, and Late Bloomers) and with age and sex as covariates. Model results are presented in Table 3 and Figure 1. Both Typical Talkers (*M*shape *=* .69, SDshape = .13) and Late Bloomers (*M*shape *=* .67, SDshape = .33) had a significantly higher proportion of shape-based words in their Time 1 vocabularies compared to Persisting Late Talkers (*M*shape *=* .54, SDshape = .39).

**Table 3.** Results of regression analyses comparing the composition of children’s Time 1 noun vocabularies based on their group. Effect sizes (*f*2) for significant effects were calculated by comparing the R2 of models with and without the overall main effect of group. Family wise alpha = .017. \* indicates significant effect.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Predictor** | ***B*** | ***se*** | ***t*** | ***p*** | ***f2*** |
| **Shape-based nouns** | Group | Persisting Late Talkers v. Typical Talkers | .52 | .05 | 10.54 | <.00001\* | .13 |
| Persisting Late Talkers v. Late Bloomers | .32 | .06 | 5.61 | <.00001\* |
| Sex | -.03 | .03 | -.98 | .330 |  |
| Age at Time 1 | .04 | .009 | 4.15 | <.00001 |  |
| **Material-based nouns** | Group | Persisting Late Talkers v. Typical Talkers | .63 | .06 | 11.07 | <.00001\* | .23 |
| Persisting Late Talkers v. Late Bloomers | .04 | .06 | .64 | .52 |
| Sex | .01 | .04 | .36 | .72 |  |
| Age at Time 1 | .07 | .01 | 6.46 | <.00001 |  |
| **Multidimensional nouns** | Group | Persisting Late Talkers v. Typical Talkers | .16 | .06 | 2.45 | .015\* | .04 |
| Persisting Late Talkers v. Late Bloomers | -.12 | .07 | -1.71 | .070 |
| Sex | .03 | .04 | .85 | .395 |  |
| Age at Time 1 | .03 | .01 | 2.15 | .032 |  |

**Figure 1.** Proportion of nouns naming categories organized by similarity in shape at Time 1 (~1.5 years) in the expressive vocabularies of children in each group. Boxplots depict the range between the 25th and 75th (box length) quartiles, mean (middle line), and outliers (dots). \* indicates significant between group differences.

**Exploration of Other Noun Vocabulary Classifications.** If Persisting Late Talkers have similar vocabulary sizes but a smaller proportion of shape-based nouns relative to Late Bloomers, they must have a higher proportion of some other type of nouns in their vocabulary. Thus, we next explored other vocabulary classifications including material-based nouns and nouns with multidimensional category structure in which no one property is key to category membership (see Figure 1). In separate multiple regression models we compared the proportion of material-based nouns and the proportion of multidimensional nouns in children’s Time 1 vocabularies based on their Time 2 groups, including age and sex as covariates (Table 3). Typical Talkers (*M*material *=* .15, SDmaterial = .10), had a significantly higher proportion of material-based nouns in their vocabularies than Persisting Late Talkers (*M*material *=* .11, SDmaterial = .21). Late Bloomers (*M*material *=* .08, SDmaterial = .14), did not differ significantly from Persisting Late Talkers. Similarly, as can be seen in Table 3, while Persisting Late Talkers (*M*multi *=* .35, SDmulti = .41) differed from Typical Talkers (*M*multi *=* .16, SDmulti = .12) in the proportion of nouns with multidimensional structure in their Time 1 vocabularies, they did not differ from Late Bloomers (*M*multi *=* .24, SDmulti = .34).

***Predicting Group Outcomes***

The previous analyses demonstrate that the groups differed in their Time 1 vocabulary composition, particularly with respect to the proportion of shape-based nouns they produced. In these final analyses we examine this association further, asking whether shape-based nouns are also predictive of later language outcomes. In separate logistic regression models we predicted group (Persisting Late Talker versus Typical Talker, coded 1 and 0 respectively; and Persisting Late Talker versus Late Bloomer, coded 1 and 0 respectively) by the proportion of shape-based nouns in the Time 1 vocabulary. As can be seen in Table 4, these models revealed that children who produced fewer shape-based nouns were more likely to be Persisting Late Talkers than Typical Talkers or Late Bloomers.

**Table 4.** Results of regression analyses predicting group outcomes from children’s Time 1 shape-based noun vocabularies. Family wise alpha = .025. \* indicates significant effect.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **Predictor** | ***B*** | ***se*** | ***t*** | ***p*** |
| **Persisting Late Talkers versus Typical Talkers**  | Shape-based nouns | -1.90 | .23 | -8.00 | <.00001\* |
| Sex | -.57 | .25 | -2.32 | .021\* |
| Age at Time 1 | .22 | .05 | 4.19 | .00003\* |
| **Persisting Late Talkers versus Late Bloomers** | Shape-based nouns | -.63 | .17 | -3.73 | .0002\* |
| Sex | -.43 | .27 | -1.62 | .105 |
| Age at Time 1 | 21 | .07 | 2.77 | .006\* |

**Discussion**

The primary goal of Study 1 was to compare differences in Time 1 expressive vocabulary composition for groups of late talking children with different short-term outcomes (i.e., months later). Consistent with previous work, we found that vocabulary size alone (here object noun vocabulary size) does not reliability differentiate Persisting Late Talkers from Late Blooming toddlers who catch up to their peers (Duff, Nation, et al., 2015). Instead, we found systematic differences in the *composition* of toddlers’ vocabularies—especially their object noun vocabularies—that were associated with differences in later outcomes. Like Typical Talkers, Late Bloomers not only had more nouns in their toddler vocabularies than Persisting Late Talkers, but they also had a higher proportion of nouns that name shape-based categories. Complementary analysis predicting group outcomes from shape-based nouns revealed that children not only differed in their Time 1 shape vocabularies, but that Time 1 shape vocabularies were also predictive of group membership.

Previous longitudinal work has suggested that regularities in Typical Talkers’ early vocabulary, including saying many names for categories organized by similarity in shape, trains their attention to relevant properties of objects, making subsequent word learning easier (Smith et al., 2002). Because there are so many to-be-learned words that name categories organized by shape, learning to selectively attend to shape helps children learn more words. Thus, Late Bloomers may start with smaller than average vocabularies, but their vocabularies tend to be made up of words that will be useful in learning how to learn more words (i.e., shape-based nouns). Conversely, as toddlers, Persisting Late Talkers may be able to say fewer nouns on average and the nouns they can say may also be less useful for building their vocabularies because they are ones for which no particular property is characteristic. Thus, Persisting Late Talkers may not only start with smaller than average vocabularies, but they may also start with vocabularies that do not help them learn how to learn words—potentially slowing subsequent vocabulary growth. We return to the issue of why some children (i.e., Persisting Late Talkers) initially acquire vocabularies with different compositions in the General Discussion.

**Study 2**

In Study 2, we examined the longer-term outcomes (2-5 years later) of a subset of the participants in Study 1 for whom clinical diagnosis data were available. In this exploratory study, we asked whether toddler vocabulary composition varied for those who later received a diagnosis of Developmental Language Disorder (DLD) relative to those who did not. Because children with DLD very frequently have other speech and reading impairments (e.g., Snowling et al., 2020), any child that had a diagnosis of DLD was classified into the DLD group, regardless of other diagnoses. Any child with any other diagnosis or diagnoses (learning disability, speech impairment, or dyslexia) that did not have a DLD diagnosis was classified into the Other Diagnosis group.

**Methods**

***Participants***

Participants included 569 children (*M =* 16.00 months; 271 girls) in our Time 1 sample who had originally been in Thal and colleagues’ (2013) study. As part of their involvement in that original study, children had participated in annual visits between 4-7 years of age during which parents reported any formal diagnosis of language impairment, dyslexia, speech impairment, or learning disability. Any reports of diagnosis were then confirmed via clinical report. Thal and colleagues shared these diagnosis data with our team.

***Group Characterization***

Using diagnosis data, we categorized children as receiving a developmental language disorder diagnosis[[4]](#footnote-5) (DLD, then Specific Language Impairment or SLI), receiving another diagnosis (Other Diagnosis) at any time between 4 and 7 years of age, or not receiving a clinical diagnosis (No Diagnosis). Of the 10 children in the DLD group in this study, 2 reported only a diagnosis of DLD; 5 reported an additional diagnosis of speech impairment; 2 reported additional diagnoses of speech impairment and learning disability; and 1 reported additional diagnoses of speech impairment, learning disability, and dyslexia. Group characteristics are listed in Table 5. As can be seen in the Table, the ratio of boys to girls differs across the three groups, consistent with previous reports that boys are more likely to be diagnosed with DLD than girls (Adani & Cepanec, 2019; Tomblin et al., 1997). We therefore include sex as a covariate in our models.

**Table 5. Study 2 group characteristics**. All children in this sample were 16 months old at Time 1. Other Diagnosis includes children who received a diagnosis of dyslexia, speech impairment, or learning disability between ages 4-7 years. Vocabulary sizes are listed as *M* (SD).

|  |  |  |
| --- | --- | --- |
| **Group**  | **Sample Size** | **Vocabulary size at Time 1** |
| DLD | 10 (2 girls) | 74.60 words(88.99 words) |
| Other Diagnosis  | 34 (12 girls) | 40.32 words(32.55 words) |
| No Diagnosis | 523 (257 girls) | 59.33 words(66.99 words) |

**Results**

The goal of the analyses was to examine whether differences in vocabulary size or composition at 16 months were associated with later group membership outcomes. Vocabulary composition was measured and analyzed in the same way as in Study 1. Children receiving a DLD diagnosis were used as the reference group in the regression analyses.

***Differences in Object Noun Vocabulary Size***

We first examined differences in children’s object noun vocabulary size. Using a multiple regression model, we compared the *number* of object nouns in children’s Time 1 vocabulary based on their diagnosis group (DLD, Other Diagnosis, No Diagnosis), including sex as a covariate. Time 1 object noun vocabulary size was not associated with later language impairment. In fact, children with DLD (*M =* 43.20 words, *SD* = 56.75) had slightly *larger* Time 1 object noun vocabularies than those who received Other Diagnoses (*M =* 19.82 words, *SD* = 20.98), *B =* -24.12, se = 14.00, *t =* -1.72, *p =* .085, although not significantly so, and similarly sized vocabularies to those with No Diagnosis (*M =* 31.57 words, *SD* = 39.42), *B =* -13.04, se = 12.45, *t =* -1.05, *p =* .296. Sex was not associated with significant differences in object noun vocabulary size, *B =* 4.83, se = 3.29, *t =* 1.47, *p =* .142.

***Differences in Vocabulary Composition***

To assess differences in vocabulary composition, we first performed a chi-square analysis comparing the number of words in each classification said by each group. Although we did not find overall group differences, *X2*(4) *=* 6.99*, p=.*137, weconducted several planned follow-up analyses on this result using the same procedure as in Study 1:separate multiple regression analyses to examine whether there were specific differences in vocabulary composition across groups.

**Shape-based Noun Vocabulary.** We examined differences in children’s vocabulary composition with respect to the log transformed proportion of shape-based nouns in their Time 1 object noun vocabularies. Model results are presented in Table 6 and Figure 2. Children with DLD (*M*shape*=* .46, *SD*shape = .35) had a significantly lower proportion of shape-based nouns in their vocabularies at Time 1 than both children with No Diagnosis (*M*shape= .67, *SD*shape = .21), *p* = .002, and children who received Other Diagnoses (*M*shape= .62, *SD*shape = .26), *p* = .037. Overall, these results build on the results of Experiment 1, further suggesting group differences in the *composition* of children’s early noun vocabulary relate to later outcomes.

**Table 6.** Results of regression analyses comparing the composition of children’s Time 1 noun vocabularies based on their diagnosis group at 4-7 years. Effect sizes (*f*2) for significant effects were calculated by comparing the R2 of models with and without the overall main effect of group. Family wise alpha = .017. \* indicates significant effect.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Predictor** | ***B*** | ***se*** | ***t*** | ***p*** | ***f2*** |
| **Shape-based nouns** | Group | DLD v. No Diagnosis | .49 | .14 | 3.43 | .0006\* | .02 |
| DLD v. Other Diagnosis | .10 | .16 | 2.48 | .013\* |
| Sex | -.05 | .04 | -1.39 | 164 |  |
| **Material-based nouns** | Group | DLD v. No Diagnosis | .04 | .19 | .23 | .819 | .0003 |
| DLD v. Other Diagnosis | .006 | .22 | -.03 | .979 |
| Sex | .05 | .05 | .97 | .331 |  |
| **Multidimensional nouns** | Group | DLD v.No Diagnosis | -.20 | .19 | -1.06 | .290 | .002 |
| DLD v.Other Diagnosis | -.23 | .21 | -1.11 | .266 |
| Sex | .06 | .05 | 1.21 | .227 |  |

****

**Figure 2.** Proportion of nouns naming categories organized by similarity in shape at Time 1 (~1.5 years) in the expressive vocabularies of children in each diagnosis group. Boxplots depict the range between the 25th and 75th quartiles (box length), mean (middle line), and outliers (dots). \* indicates significant between group differences.

**Explorations of Other Noun Vocabulary Classifications.** Finally, as in Study 1, we explored whether the diagnosis groups’ Time 1 vocabularies differed with respect to the proportion of material-based nouns or the proportion of nouns naming multidimensional categories. As can be seen in Figure 2 and Table 6, children with DLD (*M*material *=* .15, *SDmaterial* = .17) did not significantly differ from children with No Diagnosis (*M*material *=* .13, *SDmaterial* = .13), or children who received Other Diagnoses (*M*material *=* .15, *SDmaterial* = .18), in the proportion of material-based nouns in their Time 1 vocabularies, *ps* > .10. Similarly, as can also be seen in Table 5, although children with DLD (*M*multi *=* .39, *SD*multi = .43) tended to have larger proportions of nouns naming multidimensional categories in their Time 1 vocabularies than those with No Diagnosis (*M*multi *=* .19, *SD*multi = .22), and those who received Other Diagnoses (*M*multi *=* .23, *SD*multi = .29), these differences were not significant.

***Predicting Group Outcomes***

The previous analyses demonstrated that the diagnosis groups differed in their Time 1 vocabulary composition, particularly with respect to the proportion of shape-based nouns they produced. In these final analyses we examine this association further, asking whether shape-based nouns are also predictive of later diagnoses. In separate logistic regression models we predicted group (DLD versus No Diagnosis, coded 1 and 0 respectively; and DLD versus Other Diagnosis, coded 1 and 0 respectively) from the proportion of shape-based nouns in the Time 1 vocabulary. As can be seen in Table 7, results of these models revealed that children who produced fewer shape-based nouns were more likely to be diagnosed with DLD than to have No Diagnosis. Shape vocabulary did not distinguish children diagnosed with DLD from children with Other Diagnoses, however.

**Table 7.** Results of regression analyses predicting diagnosis outcomes from children’s Time 1 shape-based noun vocabularies. Family wise alpha = .025. \* indicates significant effect.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Model** | **Predictor** | ***B*** | ***se*** | ***t*** | ***p*** |
| **DLD versus No Diagnosis**  | Shape-based nouns | -1.19 | .39 | -3.03 | .002\* |
| Sex | -1.49 | .81 | -1.84 | .066 |
| **DLD versus Other Diagnosis**  | Shape-based nouns | -.78 | .50 | -1.55 | .121 |
| Sex | -.69 | .89 | -.78 | .437 |

**Discussion**

The goal of Study 2 was to explore toddler vocabulary measures associated with longer-term outcomes, including clinical diagnoses. Object noun vocabulary *size* in the second year was not associated with diagnoses between 4-7 years, suggesting the specific number of words said during toddlerhood is not enough to predict long-term outcomes. However, like Study 1, we found systematic differences in vocabulary *composition*, specifically in the proportions of noun types in children’s toddler vocabularies and their developmental outcomes. Although only a small number of children in this study received a DLD diagnosis, those who did tended to have a smaller proportion of nouns naming categories organized by similarity in shape than children in the other groups. Furthermore, as in Study 1, we found that not only did groups differ in the proportion of shape-based nouns they produced, but shape-based nouns were also predictive of later diagnosis, distinguishing children with No Diagnosis from children who went on to receive a diagnosis of DLD. Together, these results support the idea that the specific words a child learns early in development create a foundation for future learning (see also Borovsky et al., 2021).

**General Discussion**

The goal of the current study was to assess whether vocabulary composition differs systematically between late talking toddlers who continue to have delays and those who catch up. In Study 1, we retrospectively analyzed the vocabularies of 850 toddlers classified as either a late talker at both an earlier and later time point (Persisting Late Talkers), classified as a late talker early in vocabulary development but not later (Late Bloomers), or classified as showing typical vocabulary development (Typical Talkers). We found that overall noun vocabulary *size* did not differ between Late Bloomers and Persisting Late Talkers. Rather, the vocabulary *composition* of Late Bloomers included a higher proportion of object nouns that name shape-based categories compared to Persisting Late Talkers. In Study 2, we focused on the subset of our sample for which a later diagnosis, or lack thereof, was known. Again, we found that toddler vocabulary composition, differentiated children with a diagnosis of DLD and typically developing children years later—children who had a lower proportion of shape-based nouns in their toddler vocabularies were more likely to receive a DLD diagnosis than to not receive any diagnosis. Together our analyses suggest that toddlers who say a smaller proportion of shape-based nouns are more likely to continue to show language delays months and years later. Although the small number of children with DLD diagnoses in our second study means that the results should be considered preliminary, this finding nevertheless points to the composition of a child’s early vocabulary as one potentially useful early indicator of risk of subsequent language development difficulties. We discuss these findings in terms of developmental cascades and the ways in which many processes come together to support word learning.

**The Role of Early Vocabulary Regularities in Language Development**

At a broad level, the current findings fit with a larger literature highlighting regularities in the language children hear and learn and how such regularities can help them “learn to learn” language. Of particular relevance is previous work suggesting that regularities in the composition of children’s early vocabularies may have a “snowball effect” that boosts later word learning: because many of the nouns English-learning children are exposed to and learn early in vocabulary development are count nouns, names for solid things, and/or names for categories well organized by similarity in shape, the attention of typically developing toddlers becomes automatically directed towards just the right property of novel objects to support future learning—shape. This boosts the rapid learning of even more of the nouns they are exposed to early (Gershkoff‐Stowe & Smith, 2004; Perry et al., 2010; Samuelson, 2002; Smith et al., 2002). Indeed, differences in vocabulary composition may also be implicated in other related disorders, such as autism spectrum disorder (Ellis Weismer et al., 2011; Haebig et al., 2021; Jiménez et al., 2021) and dyslexia (Koster et al., 2005). One implication of the current findings is that an early vocabulary that contains a low proportion of names for categories organized by shape does not support learning of concrete nouns efficiently, resulting in smaller vocabularies, less growth, and the potential for future language delays and a diagnosis of DLD.

**Differences in Attending to and Remembering Visual Object Properties**

Of course, vocabulary composition is just one marker of potential delays in vocabulary growth and is itself the product of prior developmental processes. A critical open question is why the Persisting Late Talkers and children with eventual DLD diagnoses in our sample had a lower proportion of shape-based words in their Time 1 vocabularies. This could be due to difficulties with object processing and the extraction of abstract object shape features, skills that are related to vocabulary development (Pereira & Smith, 2009; Smith, 2003; Son et al., 2008; Yee et al., 2012). Indeed, both late talkers (Jones, 2003; Perry & Kucker, 2019) and children with DLD (Collisson et al., 2014) are less likely to attend to shape when generalizing novel nouns. Relatedly, work by Borgström and colleagues shows parallels between early shape recognition at 20-months and vocabulary at 24-months (Borgström et al., 2015) and language skills at age 6 (Borgström et al., 2019). Thus, it is possible that the Persisting Late Talkers in our sample might have had fewer names for shape-based categories in their toddler vocabularies because they had early difficulties extracting abstract object shape, causing them to have a harder time learning names for categories of objects with complex shapes. It is also possible, however, that a lower proportion of shape-based words could result from enhanced attention to features other than shape, resulting in more multidimensional words in the vocabulary.

In either case, the idea that the composition of children’s initial vocabulary provides a foundation for learning-to-learn more words suggests that even a small difference in the initial composition of children’s vocabularies could cascade into delays that persist longer-term (see also Borovsky et al., 2021). A critical direction for future work, that will build on the retrospective associations we found here, will be to incorporate longitudinal measures of children’s object noun vocabulary composition, particularly the proportion of shape nouns, with general vocabulary measures to assess potential reciprocal impacts on attention to shape and vocabulary composition. The heterogeneity of late talkers, however, means that such studies would likely need to include a large number of participants and a diverse sample to enable examination of the individual differences within late talkers (c.f. Perry & Kucker, 2019). Thus, one promising avenue we are pursuing is the creation of a data repository for data from studies of the shape bias in the past 30 years. Such a repository would enable examination of the interactions between vocabulary composition and attention to shape across a substantial number of children, but also facilitate study of how these characteristics interact with other factors shown to influence early noun learning such as the syntax of the naming context or children’s manipulations of the objects.

***Influences of Domain General Cognitive Processes***

Even more generally, the differences in the vocabulary composition of Persisting Late Talkers and other children could result from deficiencies in any one of the many other processes that support language development. For example, word learning draws on multiple memory processes (Vlach, 2019; Vlach & Sandhofer, 2012; Wojcik, 2013) and deficits in working memory have long been implicated in DLD (see e.g., Montgomery et al., 2016). Differences in domain general cognitive processes such as attention and memory could lead to differences in children’s proficiency at detecting and learning regularities across words. Indeed, Graf Estes et al. (2009) found that children with language delays needed twice as much exposure to a speech stream in order to extract regularities and identify words. With respect to the current work, it is certainly possible that differences in the vocabulary compositions of the groups in our sample reflect individual differences in attention or memory limitations for some children. Difficulty remembering an association between naming events and object shape could prevent children from building a lexical foundation that would support learning more words.

However, there is continuing debate in the literature as to how best to understand the relation between these general cognitive processes operating early in vocabulary development and language delay (see e.g., Goffman & Gerken, 2020; Kamhi, 2019); a debate that reflects, in part, the mutually influential and interacting nature of these processes. Progress on this front will require further research examining the multiple processes supporting early vocabulary development and how vocabulary composition comes together with other variables to support long-term language success. For instance, measuring changes in visual attention, memory, vocabulary composition, and categorization over smaller time intervals should illuminate how multiple cascading factors come together to support children’s word learning (c.f. Samuelson, 2021).

**Limitations**

One potential limitation of the current work is the lack of a receptive vocabulary measure to complement our analyses of expressive vocabulary. Our vocabulary measure, the MBCDI, asks parents to report which words their child says, meaning the children in our sample might have understood some words that they did not yet say. Receptive and expressive language skills are generally highly correlated with respect to vocabulary size (Ring & Fenson, 2000). We chose to focus on expressive vocabulary because previous studies linking vocabulary size to object perception (Jones & Smith, 2005; Yee et al., 2012) and attention to shape when learning new nouns (Colunga & Sims, 2017; Samuelson & Smith, 1999), as well as mini-training studies that show relations between attention to shape during noun generalization tasks and vocabulary accelerations (Perry et al., 2010; Samuelson, 2002; Smith et al., 2002), have all used expressive, rather than receptive, vocabulary as the main measure. It is nevertheless possible that the specific nouns in a child’s expressive vocabulary were not representative of the types of nouns they understood. For example, some Persisting Late Talkers may have understood more shape-based nouns than they said. However, that sort of mismatch between receptive and expressive vocabulary composition would likely only add noise to our analyses. That we found systematic group differences in expressive vocabulary alone suggests that the number of shape-based nouns children say is one factor associated with later language outcomes.

 Another potential limitation of the study is the surprisingly small number of participants in the sample who received a diagnosis of DLD. Of the 569 participants for whom we had 16-month and 28-month MBCDI data and diagnosis data, only 10 (~2%) were reported to have a DLD diagnosis made between age 4 and 7 years. Previous work estimates DLD prevalence at 7% (Tomblin et al., 1997). It is possible that some children had DLD but were missed by the screening procedure either because their parents did not report the diagnosis to the experimenters or because they had not yet been diagnosed by the time they were 7 years old. Future research with larger numbers of children with DLD diagnoses will be needed to assess the generalizability of our exploratory Experiment 2 findings.

**Conclusions**

Late talkers are an immensely heterogeneous group including both children who later catch up to their peers but also children who continue to have persistent delays, and for some, a later DLD diagnoses. Few studies have identified factors present in toddler’s vocabularies that are robustly associated with long-term language outcomes. Here, our retrospective comparison of toddler vocabulary composition, specifically the proportion of names for categories organized by shape, helps to illuminate one early marker of risk for persisting language delays. Future work investigating vocabulary composition in combination with other cognitive and environmental factors will be an important next step for understanding the cascading processes that cause differences in toddler’s vocabulary trajectories. The current study highlights the importance of studying not just how many words children say, but also *what types of words* they say.

**References**

Adani, S., & Cepanec, M. (2019). Sex differences in early communication development: Behavioral and neurobiological indicators of more vulnerable communication system development in boys. *Croatian Medical Journal*, *60*(2), 141–149. https://doi.org/10.3325/cmj.2019.60.141

Armstrong, R., Scott, J. G., Whitehouse, A. J. O., Copland, D. A., Mcmahon, K. L., & Arnott, W. (2017). Late talkers and later language outcomes: Predicting the different language trajectories. *International Journal of Speech-Language Pathology*, *19*(3), 237–250. https://doi.org/10.1080/17549507.2017.1296191

Beckage, N., Smith, L., & Hills, T. (2011). Small Worlds and Semantic Network Growth in Typical and Late Talkers. *PLOS ONE*, *6*(5), e19348. https://doi.org/10.1371/journal.pone.0019348

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

Borgström, K., Torkildsen, J. von K., & Lindgren, M. (2015). Event-related potentials during word mapping to object shape predict toddlers’ vocabulary size. *Developmental Psychology*, *6*, 143. https://doi.org/10.3389/fpsyg.2015.00143

Borgström, K., Torkildsen, J. von K., Sahlén, B., & Lindgren, M. (2019). Brain Measures of Toddlers’ Shape Recognition Predict Language and Cognitive Skills at 6–7 Years. *Frontiers in Psychology*, *10*, 1945. https://doi.org/10.3389/fpsyg.2019.01945

Borovsky, A., Thal, D., & Leonard, L. B. (2021). Moving towards accurate and early prediction of language delay with network science and machine learning approaches. *Scientific Reports*, *11*(1), 8136. https://doi.org/10.1038/s41598-021-85982-0

Botting, N., & Conti-Ramsden, G. (2008). The role of language, social cognition, and social skill in the functional social outcomes of young adolescents with and without a history of SLI. *British Journal of Developmental Psychology*, *26*(2), 281–300. https://doi.org/10.1348/026151007X235891

Collisson, B. A., Grela, B., Spaulding, T., Rueckl, J. G., & Magnuson, J. S. (2014). Individual differences in the shape bias in preschool children with specific language impairment and typical language development: Theoretical and clinical implications. *Developmental Science*, *18*(3), 373–388. https://doi.org/10.1111/desc.12219

Colunga, E., & Sims, C. E. (2017). Not Only Size Matters: Early-talker and Late-talker Vocabularies Support Different Word-learning Biases in Babies and Networks. *Cognitive Science*, *41*(S1), 73–95. https://doi.org/10.1111/cogs.12409

Colunga, E., & Smith, L. B. (2005). From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review*, *112*(2), 347–382. https://doi.org/10.1037/0033-295X.112.2.347

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*(1), 125–127. https://doi.org/10.3758/BF03203646

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

Dollaghan, C. A. (2013). Late talker as a clinical category: A critical evaluation. In *Late talkers: Language development, interventions, and outcomes* (pp. 91–112). Brookes.

Duff, F. J., Nation, K., Plunkett, K., & Bishop, D. V. M. (2015). Early prediction of language and literacy problems: Is 18 months too early? *PeerJ*, *3*, e1098. https://doi.org/10.7717/peerj.1098

Duff, F. J., Reen, G., Plunkett, K., & Nation, K. (2015). Do infant vocabulary skills predict school-age language and literacy outcomes? *Journal of Child Psychology and Psychiatry*, *56*(8), 848–856. https://doi.org/10.1111/jcpp.12378

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

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: Essays in honor of Robin S. Chapman* (pp. 83–101). Lawrence Erlabum Associates.

Ellis Weismer, S., Gernsbacher, M. A., Stronach, S., Karasinski, C., Eernisse, E. R., Venker, C. E., & Sindberg, H. (2011). Lexical and Grammatical Skills in Toddlers on the Autism Spectrum Compared to Late Talking Toddlers. *Journal of Autism and Developmental Disorders*, *41*(8), 1065–1075. https://doi.org/10.1007/s10803-010-1134-4

Ellis Weismer, S., Venker, C. E., Evans, J. L., & Moyle, M. J. (2013). Fast mapping in late-talking toddlers. *Applied Psycholinguistics*, *34*(1), 69–89. https://doi.org/10.1017/S0142716411000610

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., & Stiles, J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development*, *59*(5), i–185. https://doi.org/10.2307/1166093

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

Fernald, A., Perfors, A., & Marchman, V. A. (2006). Picking up speed in understanding: Speech processing efficiency and vocabulary growth across the 2nd year. *Developmental Psychology*, *42*(1), 98–116. https://doi.org/10.1037/0012-1649.42.1.98

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2017). Wordbank: An open repository for developmental vocabulary data\*. *Journal of Child Language*, *44*(3), 677–694. https://doi.org/10.1017/S0305000916000209

Gershkoff‐Stowe, L., & Smith, L. B. (2004). Shape and the First Hundred Nouns. *Child Development*, *75*(4), 1098–1114. https://doi.org/10.1111/j.1467-8624.2004.00728.x

Goffman, L., & Gerken, L. (2020). An alternative to the procedural∼declarative memory account of developmental language disorder. *Journal of Communication Disorders*, *83*, 105946. https://doi.org/10.1016/j.jcomdis.2019.105946

Graf Estes, K. (2009). From Tracking Statistics to Learning words: Statistical Learning and Lexical Acquisition. *Language and Linguistics Compass*, *3*(6), 1379–1389. https://doi.org/10.1111/j.1749-818X.2009.00164.x

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science*, *18*(3), 254–260. https://doi.org/10.1111/j.1467-9280.2007.01885.x

Haebig, E., Jiménez, E., Cox, C. R., & Hills, T. T. (2021). Characterizing the early vocabulary profiles of preverbal and minimally verbal children with autism spectrum disorder. *Autism*, *25*(4), 958–970. https://doi.org/10.1177/1362361320973799

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

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology*, *27*(2), 236–248.

Jiménez, E., Haebig, E., & Hills, T. T. (2021). 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*, *51*(9), 3109–3125. https://doi.org/10.1007/s10803-020-04772-1

Jiménez, E., & Hills, T. (2017). Network analysis of a large sample of typical and late talkers. *Proceedings of the Annual Cognitive Science Society Conference*.

Johnson, C. J., Beitchman, J. H., & Brownlie, E. B. (2010). Twenty-Year Follow-Up of Children With and Without Speech-Language Impairments: Family, Educational, Occupational, and Quality of Life Outcomes. *American Journal of Speech-Language Pathology*, *19*(1), 51–65. https://doi.org/10.1044/1058-0360(2009/08-0083)

Jones, S. S. (2003). Late Talkers Show No Shape Bias in a Novel Name Extension Task. *Developmental Science*, *6*(5), 477–483. https://doi.org/10.1111/1467-7687.00304

Jones, S. S., & Smith, L. B. (2005). Object name learning and object perception: A deficit in late talkers. *Journal of Child Language*, *32*(1), 223–240. https://doi.org/10.1017/S0305000904006646

Kamhi, A. G. (2019). Speech-language development as proceduralization and skill learning: Implications for assessment and intervention. *Journal of Communication Disorders*, *82*, 105918. https://doi.org/10.1016/j.jcomdis.2019.105918

Koster, C., Been, P. H., Krikhaar, E. M., Zwarts, F., Diepstra, H. D., & Van Leeuwen, T. H. (2005). Differences at 17 months: Productive language patterns in infants at familial risk for dyslexia and typically developing infants. *Journal of Speech, Language, and Hearing Research*, *48*(2).

Kucker, S. C., & Seidler, E. (2022). The timescales of word learning in children with language delays: In-the-moment mapping, retention, and generalization. *Journal of Child Language*, 1–29. https://doi.org/10.1017/S0305000921000817

Landau, B., Smith, L. B., & Jones, S. (1992). Syntactic context and the shape bias in children’s and adults’ lexical learning. *Journal of Memory and Language*, *31*(6), 807–825. https://doi.org/10.1016/0749-596X(92)90040-5

Le, H. N. D., Le, L. K. D., Nguyen, P. K., Mudiyanselage, S. B., Eadie, P., Mensah, F., Sciberras, E., & Gold, L. (2020). Health-related quality of life, service utilization and costs of low language: A systematic review. *International Journal of Language & Communication Disorders*, *55*(1), 3–25. https://doi.org/10.1111/1460-6984.12503

Longobardi, E., Spataro, P., Frigerio, A., & Rescorla, L. (2016). Language and social competence in typically developing children and late talkers between 18 and 35 months of age. *Early Child Development and Care*, *186*(3), 436–452. https://doi.org/10.1080/03004430.2015.1039529

Macroy-Higgins, M., & Montemarano, E. A. (2016). Attention and word learning in toddlers who are late talkers. *Journal of Child Language*, *43*(5), 1020–1037. https://doi.org/10.1017/S0305000915000379

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

Manhardt, J., & Rescorla, L. (2002). Oral narrative skills of late talkers at ages 8 and 9. *Applied Psycholinguistics*, *23*(1), 1–21. https://doi.org/10.1017/S0142716402000012

McGregor, K. K. (2020). How We Fail Children With Developmental Language Disorder. *Language, Speech, and Hearing Services in Schools*, *51*(4), 981–992. https://doi.org/10.1044/2020\_LSHSS-20-00003

McGregor, K. K., Goffman, L., Van, H. A. O., Hogan, T. P., & Finestack, L. H. (2020). Developmental Language Disorder: Applications for Advocacy, Research, and Clinical Service. *Perspectives of the ASHA Special Interest Groups*, *5*(1), 38–46. https://doi.org/10.1044/2019\_PERSP-19-00083

Montgomery, J. W., Evans, J. L., Gillam, R. B., Sergeev, A. V., & Finney, M. C. (2016). “Whatdunit?” Developmental changes in children’s syntactically based sentence interpretation abilities and sensitivity to word order. *Applied Psycholinguistics*, *37*(6), 1281–1309. https://doi.org/10.1017/S0142716415000570

Pereira, A. F., & Smith, L. B. (2009). Developmental changes in visual object recognition between 18 and 24 months of age. *Developmental Science*, *12*(1), 67–80. https://doi.org/10.1111/j.1467-7687.2008.00747.x

Perry, L. K., Axelsson, E. L., & Horst, J. S. (2016). Learning What to Remember: Vocabulary Knowledge and Children’s Memory for Object Names and Features. *Infant and Child Development*, *25*(4), 247–258. https://doi.org/10.1002/icd.1933

Perry, L. K., & Kucker, S. C. (2019). The Heterogeneity of Word Learning Biases in Late Talking Children. *Journal of Speech, Language, and Hearing Research*, *62*(3), 554–563.

Perry, L. K., Meltzer, A. L., & Kucker, S. C. (2021). Vocabulary Development and the Shape Bias in Children With Hearing Loss. *Journal of Speech, Language, and Hearing Research*. https://doi.org/10.1044/2021\_JSLHR-21-00003

Perry, L. K., & Saffran, J. R. (2017). Is a pink cow still a cow? Individual differences in toddlers’ vocabulary knowledge and lexical representations. *Cognitive Science*, *41*(4), 1090–1105. https://doi.org/10.1111/cogs.12370

Perry, L. K., & Samuelson, L. K. (2011). The shape of the vocabulary predicts the shape of the bias. *Frontiers in Psychology*, *2*. https://doi.org/10.3389/fpsyg.2011.00345

Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn Locally, Think Globally: Exemplar Variability Supports Higher-Order Generalization and Word Learning. *Psychological Science*, *21*(12), 1894–1902. https://doi.org/10.1177/0956797610389189

Potrzeba, E. R., Fein, D., & Naigles, L. (2015). Investigating the Shape Bias in Typically Developing Children and Children with Autism Spectrum Disorders. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00446

R Core Team. (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. http://www.R-project.org/

Rescorla, L. (2000). Do late-talking toddlers turn out to have reading difficulties a decade later? *Annals of Dyslexia*, *50*(1), 85–102. https://doi.org/10.1007/s11881-000-0018-2

Rescorla, L. (2009). Age 17 Language and Reading Outcomes in Late-Talking Toddlers: Support for a Dimensional Perspective on Language Delay. *Journal of Speech, Language, and Hearing Research*, *52*(1), 16–30. https://doi.org/10.1044/1092-4388(2008/07-0171)

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

Rescorla, L., Roberts Julie, & Dahlsgaard Katherine. (1997). Late Talkers at 2. *Journal of Speech, Language, and Hearing Research*, *40*(3), 556–566. https://doi.org/10.1044/jslhr.4003.556

Ring, E. D., & Fenson, L. (2000). The correspondence between parent report and child performance for receptive and expressive vocabulary beyond infancy. *First Language*, *20*(59), 141–159. https://doi.org/10.1177/014272370002005902

Rudolph, J. M. (2017). Case History Risk Factors for Specific Language Impairment: A Systematic Review and Meta-Analysis. *American Journal of Speech-Language Pathology*, *26*(3), 991–1010. https://doi.org/10.1044/2016\_AJSLP-15-0181

Rudolph, J. M., & Leonard, L. B. (2016). Early Language Milestones and Specific Language Impairment. *Journal of Early Intervention*, *38*(1), 41–58. https://doi.org/10.1177/1053815116633861

Samuelson, L. K. (2002). Statistical Regularities in Vocabulary Guide Language Acquisition in Connectionist Models and 15-20-month-olds. *Developmental Psychology*, *38*(6), 1016–1037. https://doi.org/10.1037/0012-1649.38.6.1016

Samuelson, L. K. (2021). Toward a Precision Science of Word Learning: Understanding Individual Vocabulary Pathways. *Child Development Perspectives*, *15*(2), 117–124. https://doi.org/10.1111/cdep.12408

Samuelson, L. K., & Smith, L. B. (1999). Early noun vocabularies: Do ontology, category structure and syntax correspond? *Cognition*, *73*(1), 1–33. https://doi.org/10.1016/S0010-0277(99)00034-7

Samuelson, L. K., & Smith, L. B. (2000). Children’s Attention to Rigid and Deformable Shape in Naming and Non-Naming Tasks. *Child Development*, *71*(6), 1555–1570. https://doi.org/10.1111/1467-8624.00248

Singleton, N. C. (2018). Late Talkers: Why the Wait-and-See Approach Is Outdated. *Pediatric Clinics*, *65*(1), 13–29. https://doi.org/10.1016/j.pcl.2017.08.018

Slone, L. K., & Sandhofer, C. M. (2017). Consider the Category: The Effect of Spacing Depends on Individual Learning Histories. *Journal of Experimental Child Psychology*, *159*, 34–49. https://doi.org/10.1016/j.jecp.2017.01.010

Smith, L. B. (2003). Learning to Recognize Objects. *Psychological Science*, *14*(3), 244–250. https://doi.org/10.1111/1467-9280.03439

Smith, L. B. (2009). From Fragments to Geometric Shape: Changes in Visual Object Recognition Between 18 and 24 Months. *Current Directions in Psychological Science*, *18*(5), 290–294. https://doi.org/10.1111/j.1467-8721.2009.01654.x

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name Learning Provides On-the-Job Training for Attention. *Psychological Science*, *13*(1), 13–19. https://doi.org/10.1111/1467-9280.00403

Smith, L. B., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568. https://doi.org/10.1016/j.cognition.2007.06.010

Snowling, M. J., Hayiou-Thomas, M. E., Nash, H. M., & Hulme, C. (2020). Dyslexia and Developmental Language Disorder: Comorbid disorders with distinct effects on reading comprehension. *Journal of Child Psychology and Psychiatry*, *61*(6), 672–680. https://doi.org/10.1111/jcpp.13140

Son, J. Y., Smith, L. B., & Goldstone, R. L. (2008). Simplicity and generalization: Short-cutting abstraction in children’s object categorizations. *Cognition*, *108*(3), 626–638. https://doi.org/10.1016/j.cognition.2008.05.002

Tek, S., Jaffery, G., Fein, D., & Naigles, L. R. (2008). Do Children with Autism Spectrum Disorders Show a Shape Bias in Word Learning? *Autism Research*, *1*(4), 208–222. https://doi.org/10.1002/aur.38

Thal, D. J., Marchman, V. A., & Tomblin, J. B. (2013). Late-Talking Toddlers: Characterization and Prediction of Continued Delay. In *Late Talkers: Language Development, Interventions, and Outcomes*. Paul H Brookes Publishing.

Tomblin, J. B. (2019). Developmental Language Disorder. In *International Handbook of Language Acquisition*. Routledge.

Tomblin, J. B., Records Nancy L., Buckwalter Paula, Zhang Xuyang, Smith Elaine, & O’Brien Marlea. (1997). Prevalence of Specific Language Impairment in Kindergarten Children. *Journal of Speech, Language, and Hearing Research*, *40*(6), 1245–1260. https://doi.org/10.1044/jslhr.4006.1245

Vlach, H. A. (2019). Learning to Remember Words: Memory Constraints as Double-Edged Sword Mechanisms of Language Development. *Child Development Perspectives*, *13*(3), 159–165. https://doi.org/10.1111/cdep.12337

Vlach, H. A., & Sandhofer, C. M. (2012). Fast Mapping Across Time: Memory Processes Support Children’s Retention of Learned Words. *Frontiers in Psychology*, *0*. https://doi.org/10.3389/fpsyg.2012.00046

Weber, J., & Colunga, E. (2021). *A Longitudinal Analysis of Vocabulary Structure in Persistent Late-Talkers, Late-Bloomers, and Typically-Developing Toddlers*. Society for Research In Child Development, virtual.

Wojcik, E. H. (2013). Remembering New Words: Integrating Early Memory Development into Word Learning. *Frontiers in Psychology*, *0*. https://doi.org/10.3389/fpsyg.2013.00151

Yee, M., Jones, S. S., & Smith, L. B. (2012). Changes in Visual Object Recognition Precede the Shape Bias in Early Noun Learning. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00533

Zubrick, S. R., Taylor, C. L., Rice, M. L., & Slegers, D. S. (2007). Late Language Emergence at 24 Months: An Epidemiological Study of Prevalence, Predictors, and Covariates. *Journal of Speech, Language, and Hearing Research*, *50*(6), 1562–1592. https://doi.org/10.1044/1092-4388(2007/106)

1. Although some of the work reviewed here may have used the term “Specific Language Impairment” or SLI to describe this population, we generally use the term Developmental Language Disorder throughout the manuscript to be consistent with current usage (see e.g., Bishop, 2017). [↑](#footnote-ref-2)
2. Other exploratory analyses of vocabulary composition related to lexical class can be found in the Supplemental Information. [↑](#footnote-ref-3)
3. In this paper, because our primary focus was on shape-based nouns, we combined the small number of words in vocabulary classifications that have been previously called “for material side” (n=24) that have an overlap between material organization and nonsolidity and/or mass syntax with those previously called “against the system” (n=26) that have an overlap between material organization and solidity and/or or count syntax (Perry & Samuelson, 2011; Slone & Sandhofer, 2017) in our material-based nouns classification. [↑](#footnote-ref-4)
4. At the time of these children’s diagnoses, the most commonly used term was Language Impairment (LI) or Specific Language Impairment (SLI). The research group initially reporting diagnostic information (Thal et al., 2013) now use the term Developmental Language Disorder (e.g., Borovsky et al., 2021; Tomblin, 2019). We use DLD here as well. [↑](#footnote-ref-5)