ASHA American Speech-Language-Hearing Association

JSLHR

Research Article

From Recognizing Known Words to Learning New Ones: Comparing Online Speech Processing in Typically Developing and Late-Talking 2-Year-Olds

Alexander LaTourrette,[a] Sandra Waxman,[b,c,d] Lauren S. Wakschlag,[d,e] Elizabeth S. Norton,[d,e,f] and Adriana Weisleder[d,f]

[a] Department of Psychology, University of Pennsylvania, Philadelphia [b] Department of Psychology, Northwestern University, Evanston, IL [c] Institute for Policy Research, Northwestern University, Evanston, IL [d] Institute for Innovations in Developmental Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL [e] Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL [f] Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL

ARTICLE INFO

Article History:
Received October 6, 2022
Revision received December 17, 2022
Accepted January 16, 2023

Editor-in-Chief: Stephen M. Camarata
Editor: Susan Nittrouer

https://doi.org/10.1044/2023_JSLHR-22-00580

ABSTRACT

Purpose: This study examines online speech processing in typically developing and late-talking 2-year-old children, comparing both groups' word recognition, word prediction, and word learning.

Method: English-acquiring U.S. children, from the "When to Worry" study of language and social–emotional development, were identified as typical talkers (n = 67, M_{age} = 27.0 months, SD = 1.4; Study 1) or late talkers (n = 30, M_{age} = 27.0 months, SD = 2.0; Study 2). Children completed an eye-tracking task assessing their ability to recognize both nouns and verbs, to use verbs to predict an upcoming noun's referent, and to use verbs to infer the meaning of novel nouns.

Results: Both typical and late talkers recognized nouns and verbs and used familiar verbs to predict the referents of upcoming nouns, whether the noun was familiar ("You can eat the apple") or novel ("You can eat the dax"). Late talkers were slower in using familiar nouns to orient to the target and were both slower and less accurate in using familiar verbs to identify the upcoming noun's referent. Notably, however, both groups learned and retained novel word meanings with similar success.

Conclusions: Late talkers demonstrated slower lexical processing, especially for verbs. Yet, their success in using familiar verbs to learn novel nouns suggests that, as a group, their slower processing did not impair word learning in this task. This sets the foundation for future work investigating whether these measures predict later language outcomes and can differentiate late talkers with transient delays from those with language disorders.

Around their second birthday, children typically enter a period of dramatic language growth. They learn new words at an accelerating pace (Fenson et al., 1994), diversifying their vocabularies to include increasing numbers of nouns, verbs, and adjectives (E. Bates et al., 1994) and enriching their semantic networks (Arias-Trejo & Plunkett, 2013; Borovsky et al., 2016b; Peters & Borovsky, 2019). Furthermore, they begin to combine multiple words and take advantage of ever subtler syntactic cues in the speech they hear (for a review, see Fisher et al., 2020). These advances in lexical and grammatical knowledge are also reflected in children's efficiency in spoken language comprehension. Between their second and third birthdays, children get faster at recognizing familiar nouns and verbs in fluent speech and using them to anticipate what comes next as the sentence unfolds (Fernald et al., 2006; Mani & Huettig, 2012).

Importantly, the pace of these advances is not uniform across children. In particular, some 2-year-olds, known as "late talkers," exhibit delays in expressive language, even in

Correspondence to Alexander LaTourrette: alatour@sas.upenn.edu.
Disclosure: The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.

the absence of other known neurodevelopmental or sensory disorders (Rescorla, 1989). Late talkers are often characterized by having small expressive vocabularies for their age and/or no word combinations at 2 years of age, and they are estimated to make up between 10% and 20% of the population (Reilly et al., 2007; Zubrick et al., 2007). Although being a late talker is a risk factor for future language disorders (Rescorla & Dale, 2013), the relationship between late-talker status and subsequent diagnosis is far from straightforward. Approximately 20%–40% of late talkers receive a diagnosis of developmental language disorder (DLD; Dale et al., 2003; Paul et al., 1997; Reilly et al., 2010). On the other hand, most children identified as late talkers on or before their second birthdays achieve typical vocabulary and language levels in preschool (Dale et al., 2003; Henrichs et al., 2011; Westerlund et al., 2006) or elementary school (see Rescorla, 2011, for a review). On average, however, late-talking toddlers continue to score lower than their peers on a variety of language and reading assessments into adolescence, even if they do not develop frank language disorders (Moyle et al., 2007; Rescorla, 2002, 2009; Rice et al., 2008).

Given the varied language development trajectories among late talkers and the challenge of distinguishing between these trajectories at a young age, it is necessary to go beyond broad patterns to specify the particular language skills that differentiate remitted versus persistent language delays. This is crucial for earlier identification and targeting of those young children at highest risk for persistent difficulties. Although late talkers have been historically defined by delays in their expressive, not receptive, vocabularies, some research suggests that toddlers with delays in both expressive and receptive vocabularies are most likely to experience future language delays (Ellis & Thal, 2008).

In addition, late talkers, on average, also exhibit impairments in their processing and use of known words during real-time language comprehension and word learning (Fernald & Marchman, 2012). These impairments, in turn, could lead to consistently slower vocabulary growth. As early as 17 months of age, children who are faster to process familiar words are also more successful in learning novel words, suggesting that faster lexical processing allows children to better encode novel word forms and/or their meanings (Lany, 2018). As children learn more words and become better at processing them, they can also use these familiar words to infer the meanings of novel words embedded in a sentence (Ferguson et al., 2014, 2018). Children's vocabulary size is also associated with a better understanding of the phonotactic patterns of their language, facilitating subsequent word learning (Graf et al., 2011). Moreover, longitudinal studies of processing speed and vocabulary suggest that faster processing speed predicts subsequent vocabulary growth, controlling for concurrent vocabulary size, with these relationships proving especially important for children with smaller initial vocabularies (Fernald & Marchman, 2012; Fernald et al., 2006; Peter et al., 2019). Similarly, processing speed predicts subsequent accelerated growth in the syntactic complexity of children's speech (Peter et al., 2019). Thus, late talkers' language processing abilities may play an important role in facilitating or limiting their subsequent vocabulary growth and syntactic knowledge. If this is the case, then early assessments of online word recognition and word learning abilities might help identify those late talkers at greatest risk for future language impairments. To address this possibility, this study seeks to shed light on the relationship between late talking, online word recognition, and word learning in toddlers.

Late Talkers' Vocabularies

Much of the evidence on late talkers' language skills in the second and third years of life has focused on vocabulary size and composition, with recent work suggesting that verbs represent a particularly crucial difference between late and typical talkers' lexicons. Caregiver reports indicate a lower proportion of verbs in the vocabularies of late talkers at 24 months of age than in those of typical talkers (Hadley, 2006; MacRoy-Higgins et al., 2016), although this difference does not appear to be tied to any specific class of verbs (Horvath et al., 2019). Verb learning difficulties may foreshadow later difficulties with grammar, as verb lexicon size is a strong predictor of grammatical development. In particular, spontaneous verb production at 24 months of age accounts for nearly half of the variance at 30 months of age on the Index of Productive Syntax (IPSyn), a measure of grammatical competence (Hadley et al., 2016). Late talkers also continue to score lower on the IPSyn at 3 and 4 years of age (Rescorla et al., 1997, 2000), even after a majority are within typical ranges on vocabulary measures.

These differences appear to persist even later into childhood. Children who were late talkers at the age of 2 years perform less accurately than those who were typical talkers on morphosyntactic and syntactic measures at the age of 7 years (Rice et al., 2008). Furthermore, 5- to 8-year-old children diagnosed with DLD have more difficulty than their typically developing peers in comprehending and producing verbs (Andreu, Sanz-Torrent, & Guàrdia-Olmos, 2012; Andreu, Sanz-Torrent, Legaz, & MacWhinney, 2012). This is in line with a larger literature suggesting that children who experience language delays and those with language disorders show more pronounced difficulties with verbs than with nouns (for a review, see Verhoeven & van Balkom, 2003).

Indeed, early impairments in verb comprehension may cause broader difficulties for vocabulary and syntactic development as well: Young children depend on the words they know to interpret novel syntactic constructions and infer the meanings of new words (Babineau et al., 2021). For instance, when a known verb takes a novel noun as an argument (e.g., "The dax is dancing" or "She drives the vep"), 2-year-old typical talkers can use the verb's selectional restrictions to identify the referent (Ferguson et al., 2014, 2018; Goodman et al., 1998). If late talkers know fewer verbs or their verb knowledge is less robust, efficient, or accessible, these inferences will be less frequent or effective, and words could take longer to learn.

Late Talkers' Online Language Processing

Late talkers also differ from typical talkers in the efficiency with which they process words in fluent speech. Fernald and Marchman (2012) used a looking-while-listening paradigm to assess how efficiently 18-month-old late talkers processed familiar words. Children listened to sentences with familiar nouns (e.g., "Look at the *kitty*") while looking at two images depicting the target and a distracter. Analyses of children's eye movements provided a measure of their efficiency in spoken word recognition. Children identified as late talkers at 18 months of age were slower and less accurate in identifying the referents of familiar nouns than their age-matched peers. Processing efficiency also predicted subsequent vocabulary growth: Children with faster and more accurate language processing at 18 months of age showed more rapid vocabulary growth and had larger vocabularies at 30 months of age. Critically, late talkers with higher processing efficiency at 18 months of age were more likely to catch up with typical peers in vocabulary by 30 months of age. Thus, lower processing efficiency may present an obstacle to efficient vocabulary growth.

Recent work has expanded these paradigms beyond familiar nouns to test young children's processing of verbs. Some of this work has suggested that the association between measures of verb processing efficiency and vocabulary size evident in late talkers is relatively weak or absent in typically developing 2- and 3-year-olds (Koenig et al., 2020; Valleau et al., 2018). However, these findings may be the result of methodological differences. In these studies, verb processing efficiency was assessed by measuring children's looking to pairs of dynamic events, which may complicate the use of standard eye gaze measures such as latency to look to the target. Other studies have circumvented this problem by testing children's efficiency in using a semantically informative verb to predict an upcoming noun. In this paradigm, children view static images of, for example, a ball and a cake, while hearing a sentence such as, "The boy eats the big cake." If children can process verbs rapidly in online speech, they should look more to the cake than to the ball after hearing the verb "eats," even before hearing the noun "cake." Indeed, Mani and Huettig (2012) found that 2-year-old typically developing children fixated on the semantically related image (e.g., the cake) soon after hearing the informative verbs. They also found that children's skill in anticipating the target from the semantically constraining verb was correlated with their productive vocabulary size, suggesting a relationship between verb processing and vocabulary size in typically developing children.

A recent study using this paradigm found successful verb-based prediction for 3- to 4-year-olds, both those who were developing typically and those who were suspected of having DLD (van Alphen et al., 2021). The two groups differed, however, in their verb processing efficiency: Children suspected of having DLD were slower than their typically developing peers to shift to the target image after hearing the verb. A similar pattern emerged in children's noun-based processing efficiency: When children heard a noun presented after an uninformative verb (e.g., "Look, a book!"), children suspected of having DLD were slower than their typically developing peers to shift their gaze to the target referent. Thus, both verb- and noun-based processing efficiency are associated with language ability in the preschool years. It remains unclear, however, whether differences in online processing might be evident even earlier for late-talking toddlers. We address this in the experiments below using a paradigm designed to investigate noun- and verb-based processing efficiency in 2-year-olds.

Word Learning in Late Talkers

Researchers seeking to identify the sources of late talkers' smaller productive vocabularies have also examined their ability to learn new words. Word learning paradigms provide a strong test of the different mechanisms that support or limit vocabulary acquisition and permit us to assess whether these mechanisms differ for late and typical talkers.

Studies have found mixed results on whether late talkers differ from typical talkers in word learning. Jones (2003) found that 2-year-old late talkers required a similar number of trials to learn a novel word in a habituation paradigm as typical talkers; the two groups were also similarly successful in extending the word to highly similar objects. However, the two groups differed in their willingness to generalize the word to other objects with similar shapes but different textures. This finding converges with other work suggesting more variation in the generalization biases of late talkers than in those of typical talkers, with fewer late talkers showing a reliable shape bias in noun generalization (Colunga & Sims, 2017; Perry & Kucker, 2019).

Other work has largely focused on fast-mapping procedures, in which children are briefly introduced to a series of novel objects and labels and then tested on their comprehension or production of the labels. Ellis Weismer et al. (2013) presented 30-month-old late and typical talkers with a fast-mapping test in which they learned two novel words. Although late talkers successfully learned words in this task, they were less accurate in both producing and comprehending them than typical talkers. Moreover, children's accuracy in producing the novel words was associated with their productive vocabulary; comprehension of the words likewise predicted language comprehension more generally. In a more extended word learning paradigm, MacRoy-Higgins and Montemarano (2016) taught typically developing and late-talking 2-year-olds 12 novel words over the course of 10 play sessions. Late-talking children learned significantly fewer words than their typically developing peers, although variability among the late talkers was high. Notably, late talkers attended to the objects less than typical talkers did during learning, suggesting that late talkers may have difficulty encoding the new words and their meanings, not simply in retrieving or producing them later.

In contrast, in a simpler fast-mapping task, 18-month-old typical and late talkers did not differ in their word learning outcomes (Ellis et al., 2015). Infants in this task were taught novel words for two novel objects, shown one at a time, and each object was labeled 14 times during 40 s of exposure to the object. Infants then completed comprehension tests for each word. Both late and typical talkers had difficulty learning the words, and there were no group differences in the accuracy of infants' looks or in more fine-grained measures such as their latency to fixate on the target. In another fast-mapping comprehension task featuring verbs, typically developing and late-talking 2-year-olds were similarly successful in learning verb–event mappings, especially when the linguistic contexts consistently included content nouns (Horvath & Arunachalam, 2021). Thus, both studies suggest more similarities than differences in late and typical talkers' word learning abilities.

Taken together, these findings paint a mixed picture. Although there appear to be some differences in typical and late talkers' word learning abilities or strategies, these differences vary across age groups and paradigms. An important limitation of the extant work, however, is that assessments of word learning have focused almost exclusively on unambiguous fast-mapping paradigms, which simply require infants to pair a novel label with a single visible object. Yet these "ideal" scenarios are uncommon in children's everyday learning environments, in which many potential referents for a new word are often present (Medina et al., 2011). In such circumstances, both typical and late talkers can succeed in learning words through repeated exposures and cross-situational word learning (Alt et al., 2014, 2020; Smith & Yu, 2008). However, the linguistic context in which the novel word appears can also be instrumental in identifying the target referent (Fisher et al., 2010; Gillette et al., 1999). By 2 years of age, typically developing toddlers readily recruit familiar verbs to constrain the meaning of novel nouns used as their arguments (Ferguson et al., 2014, 2018; Goodman et al., 1998). For example, in a sentence such as, "Mommy feeds the *ferret*," children use the familiar verb "feed" to infer that the target of the sentence (i.e., the intended referent of the novel noun) must be an animate and not an inanimate object. A critical question then is how late talkers learn words in these more naturalistic situations requiring them to integrate linguistic and referential context, especially as these situations depend on preexisting language knowledge and efficient processing, which may be impaired for late talkers.

This Study

This study was designed to investigate both real-time familiar word recognition and novel word learning in late-talking and typically talking toddlers and to measure the relation between these abilities in both groups. To do so, we designed a new looking-while-listening task to assess how rapidly 2-year-olds (a) recognize familiar nouns, (b) use familiar verbs to predict the referent of an upcoming familiar noun, and (c) use familiar verbs to identify and learn the referent of a novel noun.

First, we tested noun recognition by including trials from a standard looking-while-listening paradigm with familiar nouns (e.g., Fernald et al., 2006, 2008). Second, we assessed verb-based prediction abilities by testing how rapidly toddlers were able to use a familiar verb to predict an upcoming noun (e.g., "You can *eat* the apple"). Finally, we investigated typical and late talkers' ability to use familiar verbs to learn the meaning of novel nouns. On these trials, we introduced novel nouns in sentences with semantically constraining verbs (e.g., toddlers heard "You can eat the dax" when a novel food item and a novel clothing item were shown on the screen). Here, toddlers had to rely on the familiar verbs ("eat" or "wear") to infer the noun's referent (an item of food or clothing, respectively). We then tested toddlers' retention of these noun–object mappings when the informative verb (e.g., "eat") was replaced with a neutral one (e.g., "find") that could refer to either referent.

In Study 1, we implemented this task with 2-year-old typical talkers. Our goal was to compare their success in two different implementations of the task: a Massed version, in which children heard only one novel word (of two) per block, and an Interleaved version, in which children heard two different novel words in each block. In Study 2, we presented the easier Massed version to 2-year-old late talkers and compared their performance to that of the typical talkers from Study 1. By examining familiar word recognition and novel word learning in the same design, this paradigm permits us to examine the relations among these skills and identify differences between late and typical talkers.

Study 1

The goal of Study 1 was to establish a normative baseline with typically developing children in this task. To do so, we assessed their noun recognition, verb-based prediction, and verb-based word learning. In addition, we compared two different versions of this task, featuring either a Massed or an Interleaved word learning trial structure. In both versions, children learned two novel nouns. However, in the Interleaved version, children encountered both nouns within each block (e.g., learning "dax" on one trial and then "vep" two trials later). In contrast, in the Massed version, children heard only one novel noun in each block, with that noun repeated twice within the block and the nouns alternating across blocks (e.g., "dax" in Block 1 and "vep" in Block 2). The goal of this manipulation was purely methodological: We simply wanted to determine which trial structure provided the best opportunity for word learning in this task, so that we could then use that version with late talkers. Previous work has sometimes found that children learn words better in massed contexts, where the novel word is repeated across multiple, consecutive exposures (Schwab & Lew-Williams, 2016, 2020; Vlach & Johnson, 2013). However, in other studies, children successfully learned words in both massed and interleaved contexts (Benitez et al., 2020; Vlach & DeBrock, 2017), and interleaving ensures that children have an equal opportunity to learn both words regardless of their fatigue over blocks. Thus, we compared the Massed and Interleaved trial structures in Study 1 to determine which better supported learning in this context.

Method

Participants

All participants in Studies 1 and 2 were part of a larger longitudinal study of emergent language and mental health risk in toddlers (the "When to Worry" [W2W] study; Krok et al., 2022; Norton et al., 2022). Participants included toddlers with typical development and high irritability (the most robust developmental indicator of behavioral vulnerability to subsequent mental health problems; Wakschlag et al., 2019) and late talkers. All participants were recruited from the greater Chicago metropolitan area via pediatric practices, social media, and advertisements. To be eligible for the larger study, children were required to be full-term (gestational age of greater than 36 weeks), with no sensory/perceptual or neurodevelopmental diagnoses, monolingual English acquiring (less than 20% exposure to other languages), and between 24 and 32 months of age, with a biological parent willing to complete the study.

For this article, we included all children who participated in the eye-tracking task during their visit at 2 years old before March 2020; an additional three children participated in the eye-tracking task but contributed no usable trials. Other children in the larger study did not complete the eye-tracking task because they did not participate in the "2-year-old age" visit by March 2020 (n = 60); the eye tracker was out of order or experienced technical difficulties (n = 65); fatigue, time constraints, or other factors prevented the children from getting to the task (n = 14); or in-laboratory data collection was not possible, usually due to the COVID-19 pandemic (n = 21). The current sample (N = 97) did not significantly differ from the rest of the larger study's sample (N = 163) with respect to age, sex (43 female, 54 male), or ethnicity (12 Hispanic or Latino, 85 not Hispanic or Latino; $ps > .15$) but did differ in race (75 White, nine Black, one Asian, eight multiracial, four unreported). Specifically, the current sample included proportionally more White participants, $\chi^2 = 13.8$, $p < .001$, and fewer Black/African American participants, $\chi^2 = 9.46$, $p = .002$. The current sample also had a significantly higher income-to-needs ratio, $t(132) = 2.97$, $p = .0035$, and was more likely to have a primary caregiver with a college degree, $\chi^2 = 5.73$, $p = .017$. As most children with the opportunity to complete the current task did so, these differences between the current and the larger sample likely reflect differences in children's opportunity to visit the laboratory (with a functioning eye tracker), not differences due to the task itself.

The longitudinal W2W study includes yearly direct assessment visits (including the Mullen Scales of Early Learning [MSEL; Mullen, 1995] and eye-tracking tasks) and parent surveys. Parent surveys, interviews, and video chat visits occur between yearly visits. Parents completed the MacArthur–Bates Communicative Development Inventories (MCDI; Fenson et al., 2006) at multiple time points. For the current analyses, we used the MCDI data from the time point closest to the date the children completed their

Table 1. Participant characteristics in Studies 1 and 2.

	Study 1				Study 2	
	Typical talkers (Interleaved) *n* = 28		Typical talkers (Massed) *n* = 39		Late talkers (Massed) *n* = 30	
Measure	***M*** **(*SD*)**	**Range**	***M*** **(*SD*)**	**Range**	***M*** **(*SD*)**	**Range**
Age (months)	26.8 (1.2)[a]	25–29	27.1 (1.5)[a]	25–30	27.0 (2.0)[a]	24–32
MCDI words produced percentile	56.9 (22)[a]	18–95	57.4 (22)[a]	20–95	5.4 (6.3)[b]	1–20
MSEL receptive language percentile	70.1 (24)[a]	16–99	64.8 (23)[a]	24–99	39.5 (28)[b]	1–92
MSEL expressive language percentile	71 (25)[a]	16–99	63.5 (25)[a]	8–96	21 (24)[b]	1–86

Note. MCDI = MacArthur–Bates Communicative Development Inventories; MSEL = Mullen Scales of Early Learning.

[a,b]Different superscripts indicate groups that differ at $p < .05$, based on t tests; same superscripts indicate groups that do not differ significantly (for details, see the Method section under Study 2).

eye-tracking visit; as a result, the ages of children at the MCDI assessment ranged from 19.8 to 29.8 months (M = 24.8, SD = 2.1). Given the variability in age, we used age-based percentiles for words produced to classify children as typical or late talkers (see definitions below). Participants also completed the MSEL during their visit to the laboratory.

The participants in Study 1 were sixty-seven 2-year-olds (29 girls, 38 boys; one Asian, five Black, five Hispanic or Latino, five multiracial, 51 White) in the larger study who were classified as typical talkers because they (a) scored above the 15th percentile for vocabulary on the MCDI and (b) were reported to be combining words. See language scores and ages in Table 1. An additional child participated in the task but failed to contribute adequate looking (attending for a minimum of 750 ms during the target window; see the Analysis Strategy section) on any trial and was thus excluded. Each child participated in either the Interleaved version (n = 28) or the Massed version (n = 39) of the task. Independent two-sample Welch's t tests revealed no differences between versions in the children's age, $t(64) = 0.96$, $p = .34$, or vocabulary percentile, $t(59) = 0.10$, $p = .92$.

Task Design

The looking-while-listening task included five distinct trial types, with two trials of each type presented in each of four blocks (see Figure 1). On all trials, children viewed two images presented simultaneously. On Warm-up trials, which occurred at the start of each block, children viewed two familiar objects that were neither clothing nor food (e.g., a baby and a truck). These trials were designed to assess familiar-noun recognition and to add variety in items (e.g., Fernald & Marchman, 2012). On all remaining trial types, children viewed a food item and a clothing item. Trials varied in (a) whether the sentences included a neutral verb ("get," "see," or "find") or an informative verb ("wear" or "eat") and (b) whether the images were known (e.g., sock, apple) or novel (kimono, dragon fruit). Thus, there were four critical trial types: Neutral–Known (e.g., "You can see the apple"), Informative–Known (e.g., "You can eat the apple"), Informative–Novel (e.g., "You can eat the dax"), and Neutral–Novel (e.g., "Find the dax"). Neutral–Known trials (together with Warm-up trials) assessed familiar-noun recognition, and Informative–Known and Informative–Novel trials assessed verb-based prediction.

Figure 1. Task design. The task was composed of five trial types, described in the Method section. Bolded words indicate the first word the children could use to identify the target object for each trial. Two trial types assessed recognition of familiar nouns (Warm-up and Neutral–Known trials). Two trial types assessed verb-based prediction using familiar verbs (Informative–Known and Informative–Novel trials), and Neutral–Novel trials tested children's retention of the novel words introduced in Informative–Novel trials.

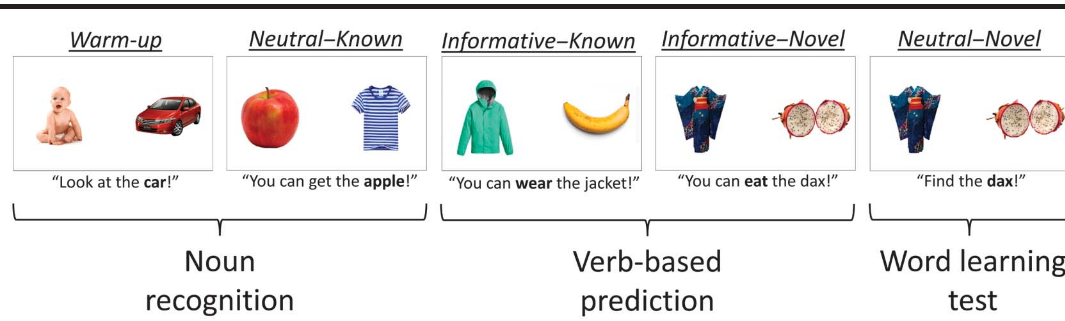
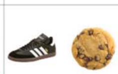


Figure 2. Sample block of trials. Each of the four blocks of trials included the same sequence of trial types, shown in the figure. There were two trials of each type (indicated as #1 and #2). Massed and Interleaved versions of the task differed only in the second Informative–Novel and Neutral–Novel trials (outlined in red). On these trials, the Massed task repeated the novel word heard earlier in the block, whereas the Interleaved task presented the other novel word. Thus, the Massed task presented only one novel word per block, with the two words presented in alternating blocks, whereas the Interleaved task presented both words in each block.

	Warm-up #1	Warm-up #2	Informative – Known #1	Informative – Known #2	Informative – Novel #1	Neutral – Novel #1	Neutral – Known #1	Informative – Novel #2	Neutral – Novel #2	Neutral – Known #2
Images:										
Interleaved Task (Audio)	Look at the car!	Look at the doggie!	You can wear the jacket!	You can eat the cookie!	You can wear the dax!	Find the dax!	You can see the shirt!	You can eat the vep!	Find the vep!	You can get the cheese!
Massed Task (Audio)	Look at the car!	Look at the doggie!	You can wear the jacket!	You can eat the cookie!	You can wear the dax!	Find the dax!	You can see the shirt!	You can wear the dax!	Find the dax!	You can get the cheese!

Informative–Novel trials also served as "teaching" trials for the novel words, allowing children to use a familiar verb to infer the referent of the novel noun. Neutral–Novel trials served as "test" trials for these novel nouns: Neutral–Novel trials always immediately followed an Informative–Novel trial and tested infants' understanding of the same novel noun heard on that previous Informative–Novel trial. The two familiar verbs "eat" and "wear" were selected because children successfully used them to predict noun meanings in previous work (e.g., Borovsky et al., 2016b; Goodman et al., 1998) and for their compatibility with multiple familiar nouns. In addition, the CHILDES (Child Language Data Exchange System) corpus suggests "eat" and "wear" are used at similarly high levels in maternal child-directed speech at 24 months of age (2,800 ppm for each verb), although children themselves use "eat" more frequently (3,400 ppm) than "wear" (870 ppm; MacWhinney, 2014; Sanchez et al., 2019). Lexical norms based on parental MCDI reports indicated that the familiar nouns were all produced by a majority of 2-year-olds (Frank et al., 2017).

Critically, consecutive Informative–Novel and Neutral–Novel trials always featured the same noun. However, these trials were presented in one of two larger task structures (see Figure 2). In the Interleaved structure, children were exposed to Informative–Novel and Neutral–Novel trials for both words within each block; in the Massed structure, children heard the same word for all Informative–Novel and Neutral–Novel trials within a block, with that word alternating across blocks (e.g., for one novel word, the learning trials would occur in Blocks 1 and 3; for the other, the learning trials would occur in Blocks 2 and 4). The order of nouns was counterbalanced across children.

Stimuli

Language stimuli consisted of 28 different target sentences, all recorded by a female native speaker of English using infant-directed speech in a soundproof booth. These stimuli corresponded to different trial types (see Figure 1). Warm-up trials consisted of eight sentences featuring familiar-object labels from a variety of categories (e.g., "Look at the baby"). Neutral–Known trials consisted of eight sentences featuring familiar-object labels ("shirt," "sock," "jacket," "shoe," "apple," "cookie," "banana," "cheese") and presented with a neutral verb with uninformative selectional restrictions (e.g., "You can get/see the apple"). Informative–Known trials consisted of eight sentences featuring the same familiar-object labels presented with an informative verb, either "wear" or "eat," as appropriate (e.g., "You can eat the apple," "You can wear the sock"). Informative–Novel (i.e., word learning) trials consisted of sentences using two novel-object labels ("dax" and "vep") with informative verbs. For each participant, one of the two novel nouns was paired with the verb "wear" (e.g., "You can wear the dax"), and the other, with the verb "eat" (e.g., "You can eat the vep"); this pairing was counterbalanced across participants. Finally, Neutral–Novel (i.e., test) trials consisted of sentences featuring the same novel-object labels as in the Informative–Novel trials but with neutral verbs (e.g., "Find the dax").

Visual stimuli consisted of 16 distinct images corresponding to the object labels above, as well as two images of novel objects: a dragon fruit and a kimono.[1] All objects

[1]Pilot testing indicated that kimonos and dragon fruit were unfamiliar to most children in our Midwestern U.S. demographic. Thus, although they are not universally novel, they do represent "novel" stimuli from the perspectives of our participants.

appeared equally often as a target and a distracter and were always presented in the same pairs. The side of the target image was counterbalanced across trials within a block.

Procedure

We used a Tobii Pro X3-120 corneal reflection eye tracker for data collection in a laboratory setting. This eye tracker has a sampling rate of 120 Hz. Stimuli were displayed on a Lenovo ThinkPad laptop with a display size of 34.5 × 19.5 cm and 1920 × 1080 screen resolution. Children were seated approximately 60 cm from the screen.

Children participated in the looking-while-listening task during a study visit in which they also completed other activities related to the larger project (not reported here). Across both task versions, 40 trials were presented in four blocks, with each block containing two trials of each type and always presented in the order illustrated in Figure 2. Before each trial in the task, an attention-getting stimulus (e.g., a multicolored spinning circle) appeared for 2–4 s in the center of the screen. On all trials, children were first presented with two object images for 2 s in silence. Then, children heard a sentence, determined by the trial type (e.g., “You can wear the shirt” for an Informative–Known trial). This was followed by a follow-up question using a neutral pronoun (e.g., “Do you like it?”) approximately 5,500 ms into the trial, to ensure children remained engaged with the objects.

Analysis Strategy

To analyze children’s performance in the task, we analyzed both (a) the proportion of time they spent looking to the target image and (b) their reaction time in looking to the target image after the onset of the disambiguating word. For Informative–Known and Informative–Novel trials, the disambiguating word was the verb; for all other trials, the disambiguating word was the noun, unless noted otherwise in the analysis. For analyses of looking proportions, we focused on toddlers’ looking within the naming window, operationalized here as looking between 233 and 1,800 ms after the disambiguating-word onset (following Fernald et al., 2008). Finally, we included only trials in which the child attended to the objects for at least 750 ms during the naming window; this yielded approximately 20 trials (SD = 9.30) per participant, with children contributing between one and 39 trials. For the children who contributed no Informative–Novel trials (n = 2 in Study 1; n = 1 in Study 2), we also excluded their Neutral–Novel trials, as they might not have had the opportunity to learn the novel words, although an analysis including these trials yielded the same conclusions.

To analyze reaction time, we included only trials in which the child was fixated on one object at word onset and then shifted to the other at least 233 ms later; this is a conservative estimate of the minimum time required to initiate a fixation in response to the word (Swingley et al., 1999). To ensure children’s gaze shifts reflected a genuine shift to fixate on the other image (i.e., children were not simply looking off-screen and returning), we also excluded any trials with a gap of more than 300 ms between the moment infants shifted off the first image and the moment they arrived at the second (cf. Fernald et al., 2008). As a result, we retained approximately 11 trials per participant for reaction time analyses.

We then fit linear mixed-effects models (D. Bates et al., 2015) with random effects of subject and, when possible, random slopes for within-subject variables (e.g., verb, block; cf. Barr et al., 2013). All categorical variables were effect-coded. The significance of effects was evaluated using Satterthwaite approximations for degrees of freedom implemented in the lmerTest package, reported rounding to the nearest integer (Kuznetsova et al., 2017; see Luke, 2017, for justification). In cases where these models failed to converge, we averaged scores by participant to conduct a Welch’s t test or analyses of variance. These tests are all noted by a footnote, as they use participants’ average scores across trials and, unlike the multilevel models, do not account for item-level variability, with a small reduction in power; however, we do not expect that this affected the results in any substantial way. All proportions were arcsine-square-root-transformed for analysis with linear models.

Results

Typical talkers attended to the task across both task versions and all trial types, providing at least 750 ms of looking during the naming window on approximately half of the trials. There was no difference in the number of trials included for children in the Interleaved, M = 21.3 trials per child, SD = 9.75, and Massed, M = 19.1 trials per child, SD = 9.0, $t(65)$ = 0.93, p = .36, versions.

Preliminary Analyses

Preliminary analyses of target looking during the naming window examined the effects of trial block (1–4), verb (“wear” vs. “eat”), and child age. Combining across trial types, there was no significant effect of block, p = .50. We did, however, observe a significant effect of verb, such that typical talkers were more successful using “eat” than “wear” to identify the target referent on the Informative–Known, $t(50)$ = 7.27, p < .0001; Informative–Novel,

$t(54) = 4.51$, $p < .0001$; and Neutral–Novel,[2] $t(40) = 3.05$, $p = .004$, trial types. However, verb did not interact with task version for any trial type, *p*s > .05. Therefore, we adopted a conservative strategy: We retained verb as a covariate in subsequent analyses of these trial types. Finally, we found no significant effects of age on performance for any trial type, *p*s > .35.

Online Word Recognition

Known Nouns

To assess noun recognition, we focused on the Warm-up and Neutral–Known trials, which feature familiar nouns and neutral verbs. On Warm-up trials, typical talkers looked more to the target object than to the other object, $M = 0.66$, $SD = 0.15$, $t(65) = 8.28$, $p < .0001$. On Neutral–Known trials, they also looked significantly more to the target object, $M = 0.72$, $SD = 0.17$, $t(64) = 8.87$, $p < .0001$. As predicted, we observed no effect of task version (Massed vs. Interleaved) for either Warm-up, $t(54) = 1.05$, $p = .30$, or Neutral–Known, $t(51) = 0.15$, $p = .88$, trials. Across both task versions, typical talkers successfully identified the referents of familiar nouns.

Known Verbs

To test whether typical talkers were also successful in using semantically informative verbs to predict the referent of an upcoming noun, we focused first on the Informative–Known trials, examining looking from 233 to 1,800 ms after the onset of the verb. Typical talkers looked to the target at above-chance rates during this window, $M = 0.66$, $SD = 0.15$, $t(63) = 7.72$, $p < .0001$. Moreover, typical talkers performed comparably in the Interleaved and Massed versions of the task, $t(57) = 1.50$, $p = .14$. There was, however, an unanticipated effect of verb, $t(50) = 7.26$, $p < .0001$. On trials including the verb "eat," typical talkers looked to the target at above-chance levels, $M = 0.77$, $SD = 0.18$, $t(61) = 10.2$, $p < .0001$. However, typical talkers' preference for the target on "wear" trials did not reach significance, $M = 0.53$, $SD = 0.21$, $t(60) = 0.89$, $p = .38$. Thus, in contrast to prior work (cf. Borovsky et al., 2016b), toddlers in the current design were significantly more successful with "eat" than with "wear."

Next, we directly assessed whether typical talkers used the semantically informative verb to predict the referent of the upcoming noun by comparing performance on Informative–Known and Neutral–Known trials in the window following the verb. As expected, typical talkers looked more to the target referent on Informative–Known trials than on Neutral–Known trials following the verb, $t(62) = 3.13$, $p = .003$. As illustrated in Figure 3 (left panel), typical talkers showed a stronger and earlier emerging preference for the target referent after hearing an informative verb (e.g., "eat") than that after hearing a neutral verb (e.g., "see"). This suggests that typical talkers successfully used the informative verb to identify the referent of the upcoming (familiar) noun.

Word Learning

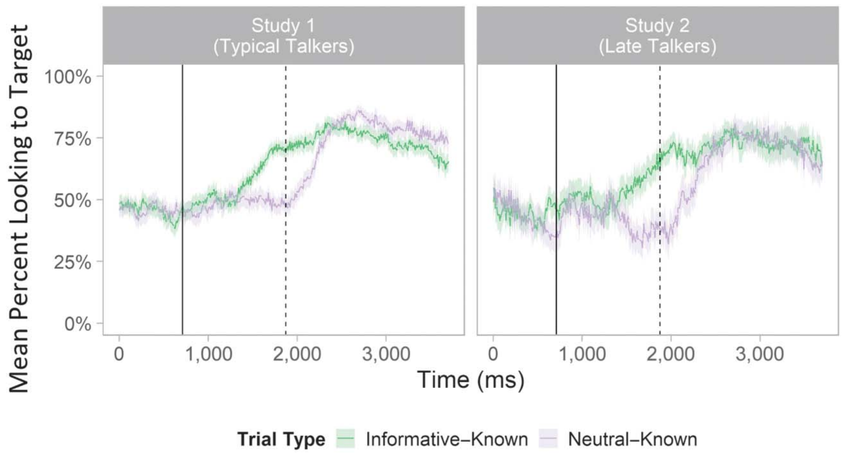
Typical talkers' successful use of a familiar verb on the Informative–Known trials set the stage to examine their use of verbs in word learning on the Informative–Novel trials. We assessed whether typical talkers also leveraged the familiar verbs to infer the referents of novel nouns (e.g., "You can eat the dax"). Children's performance on Informative–Novel trials was above chance overall, $M = 0.65$, $SD = 0.20$, $t(62) = 5.49$, $p < .0001$ (see Figure 4, blue and green bars). This suggests that 2-year-olds used informative verbs to identify the semantically appropriate referent for the novel noun. In addition, their success was not affected by the task version, $t(47) = 0.05$, $p = .96$. However, echoing the results on Informative–Known trials, there was a significant effect of verb, $t(54) = 4.51$, $p < .0001$: Typical talkers successfully identified the target on "eat" trials, $M = 0.75$, $SD = 0.20$, $t(60) = 8.62$, $p < .0001$, but did not differ from chance on "wear" trials, $M = 0.50$, $SD = 0.30$, $t(51) = 0.05$, $p = .96$.

Next, we focused on Neutral–Novel trials (e.g., "Find the dax") to assess whether typical talkers had in fact learned the meanings of the novel nouns introduced in the Informative–Novel trials. On Neutral–Novel trials, children's aggregate performance across task versions did not differ from chance, $M = 0.54$, $SD = 0.29$, $t(59) = 1.12$, $p = .26$, indicating that learning the novel label was quite difficult. Children performed numerically, though not significantly, better in the Massed version,[2] $t(46) = 1.15$, $p = .26$. While children in the Interleaved version of the task looked equally between the two images at test, $M = 0.50$, $SD = 0.23$, $t(25) = 0.14$, $p = .89$, those in the Massed task showed a preference for the target referent, which approached significance, $M = 0.57$, $SD = 0.22$, $t(33) = 1.82$, $p = .077$.

We also found a significant effect of verb on children's performance,[2] $t(39) = 3.04$, $p = .004$. As in the Informative–Novel trials, typical talkers only successfully identified the target referent for the food noun (i.e., the noun previously paired with "eat"), $M = 0.63$, $SD = 0.20$, but not for the clothing noun (i.e., the noun previously paired with "wear"), $M = 0.42$, $SD = 0.25$. There was no

[2]In this case, the most minimal mixed-effects model failed to converge or was singular, so instead, this statistic is the result of comparing participants' average scores with a linear model.

Figure 3. Time course of looking to the target on word recognition trials. Typical talkers (left panel) show earlier looking to the target referent in the Informative–Known trials (in green) than in the Neutral–Known trials (in purple), indicating that they successfully used the informative verb to identify the target of the sentence. Late talkers (right panel) show a similar pattern, looking earlier to the target on Informative–Known trials. The solid vertical line indicates the trials' average verb onset time; the dashed line indicates the average noun onset time. The *x*-axis marks the time elapsed from the trial's onset to the end of the average noun-based naming window. Colored shading represents ±1 *SEM*.



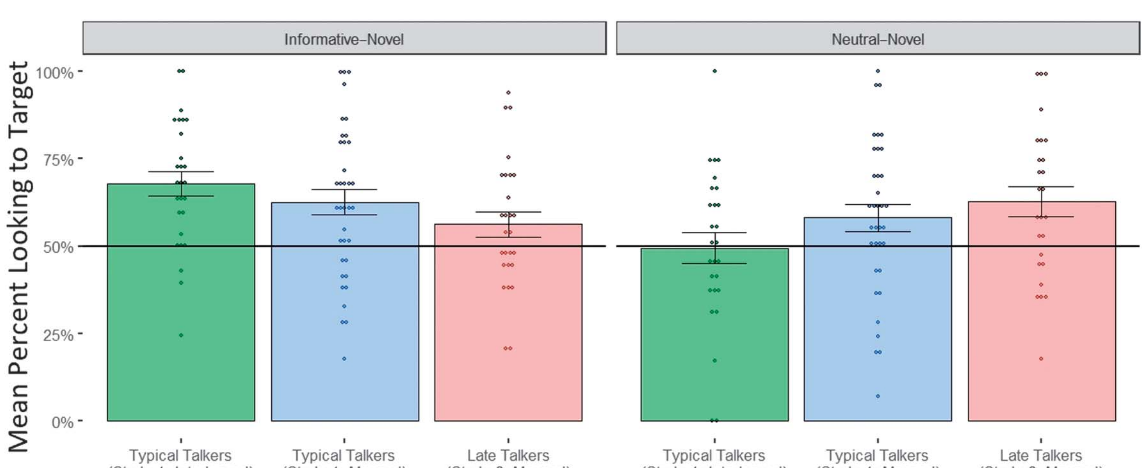
significant interaction between task version and verb,[2] $t(39) = 0.84$, $p = .41$. Overall, the verb-based word learning task was clearly quite difficult for 2-year-olds (see Figure 4), but those in the Massed version of the task did successfully learn novel words, at least when they were introduced using the verb "eat."

Processing Efficiency

We next measured how quickly typical talkers shifted their gaze from the distracter image to the target image on each trial. Specifically, we examined how quickly typical talkers shifted to the target on trials featuring a familiar *noun* as the disambiguating word (Warm-up and Neutral–Known trials) compared to those featuring a familiar informative *verb* as the disambiguating word (Informative–Known and Informative–Novel trials). Combining across trial types in this way enabled us to make the most of a sparse data set (an average of under 11 distracter-initial trials per participant, across all trial types).

Typical talkers showed rapid shifts in gaze toward the target after hearing both nouns (M = 646 ms, SD = 269) and verbs (M = 749 ms, SD = 456). Although typical

Figure 4. Proportion of time looking to the target on word learning trials. On Informative–Novel trials (left panel), typical talkers' performance did not differ between the Interleaved and Massed versions. Late talkers looked significantly less to the target referent than did the typical talkers, $p = .002$. On Neutral–Novel trials (right panel), typical talkers' performance was numerically, but not significantly, higher in the Massed version than in the Interleaved version, and no difference emerged between late and typical talkers within the Massed version.



talkers were significantly faster to shift to the target on noun trials than on verb trials, $B = 126$, $SE = 48$, $t(34) = 2.64$, $p = .012$, the reaction time difference was only around 100 ms on average. Thus, identifying an object based on a semantically informative verb appears to require more processing time than identifying an object from its noun label, but typical talkers accomplished both tasks quite rapidly.

Discussion

These findings offer converging evidence that 2-year-old typical talkers successfully use familiar verbs to anticipate the referent of an upcoming noun (Borovsky et al., 2016b; Mani & Huettig, 2012) and to infer the meaning of a novel noun (Ferguson et al., 2014). Moreover, typical talkers identified the target referent of an upcoming noun based on a semantically informative verb almost as quickly as when they heard its noun label. Thus, typical talkers efficiently processed both nouns and verbs online. Together, these results indicate that this task is sufficiently sensitive to assess toddlers' noun recognition, verb-based prediction, word learning, and processing efficiency.

The current results also offer converging evidence that massed presentation of novel words promotes word learning in children (Schwab & Lew-Williams, 2016, 2020). Although the difference between massed and interleaved conditions fell short of significance, typical talkers showed successful learning for at least one novel word in the Massed version of the task (exposed to a single novel word per block) while performing at chance in the Interleaved version (exposed to two novel words per block). This massed structure may also better resemble natural, child-directed conversation: When objects are mentioned at multiple points within a discourse, those mentions tend to involve multiple utterances clumped together (Frank et al., 2013). In Study 2, we therefore used the Massed version of this task to investigate real-time word recognition and word learning in late-talking toddlers.

Study 2

In Study 2, we examined 2-year-old late talkers' real-time word recognition and word learning. Building upon the evidence from Study 1 on word recognition and word learning abilities in typical talkers, here, we assessed late talkers' performance in the same paradigm. To provide the most sensitive test of late talkers' abilities, we presented them with the Massed trial structure from Study 1, which led to more successful word learning for typical talkers. We then compared late and typical talkers' performance across studies. For comparisons involving word learning trials, we focused specifically on the Massed version of the task. This provides a close comparison of late and typical talkers' online language processing in the same task.

Method

Participants

Participants were thirty 2-year-old children (14 girls, 16 boys; four Black, seven Hispanic or Latino, three multiracial, 16 White) who were also enrolled in the W2W study (see Table 1). All were classified as late talkers, which was defined as (a) scoring at or below the 15th percentile on the MCDI (29 participants) and/or (b) not yet combining words (10 participants). Note that, although late talkers were defined exclusively on the basis of their expressive language, late talkers as a group differed from typical talkers in both expressive, $t(56) = 8.43$, $p < .0001$, and receptive, $t(45) = 4.62$, $p < .0001$, language abilities (see Table 1). There were no significant differences between the late talkers in this study and the typical talkers in Study 1 on age, biological sex, likelihood of the mother having a college degree ($M_{\text{Late}} = 0.79$, $M_{\text{Typical}} = 0.85$), or income-to-needs ratio based on household size ($M_{\text{Late}} = 6.60$, $SD_{\text{Late}} = 10.3$; $M_{\text{Typical}} = 6.75$, $SD_{\text{Typical}} = 5.2$), all $ps > .5$.

Procedure

All materials, procedures, and data analyses were identical to the Massed version of the task in Study 1 (see Figure 2, Massed task).

Results

Late-talking children contributed an average of 17.2 trials ($SD = 9.5$) to the naming window analysis; this was not significantly different than typical talkers in Study 1, $t(95) = 1.36$, $p = .18$.

Preliminary Analyses

Preliminary analyses examined whether the proportion of looking time during the naming window was related to block, verb ("wear" vs. "eat"), or child age. As in Study 1, there was no overall effect of block, $p = .71$, but significant effects of verb emerged on Informative–Known and Informative–Novel trials, $ps < .05$, with performance higher on "eat" trials than on "wear" trials. Verb was therefore retained as a covariate in the corresponding analyses. Although no effect of verb emerged on Neutral–Novel trials, $p > .2$, we included verb as a covariate for these trials as well, in line with both the other trial types and Study 1. Finally, age was not significantly correlated with children's performance on any trial type, ps > .3.

Online Word Recognition

Known Nouns

On Warm-up trials, late talkers successfully identified the referents of familiar nouns within the naming window, just as typical talkers did in Study 1 (see Figure 5). Performance on Warm-up trials was well above chance, $M = 0.62$, $SD = 0.16$, $t(28) = 3.94$, $p = .0004$, and did not differ from that of typical talkers, $t(82) = 1.44$, $p = .16$.

On Neutral–Known trials, late talkers also looked significantly more to the target than to the distracter, $M = 0.65$, $SD = 0.20$, $t(24) = 3.68$, $p = .001$. When compared to typical talkers, this difference did not reach significance, $t(85) = 1.82$, $p = .073$. In conjunction with the Warm-up trials, this suggests that late and typical talkers show successful and largely comparable online recognition of known nouns at 2 years of age (see Figure 5).

Known Verbs

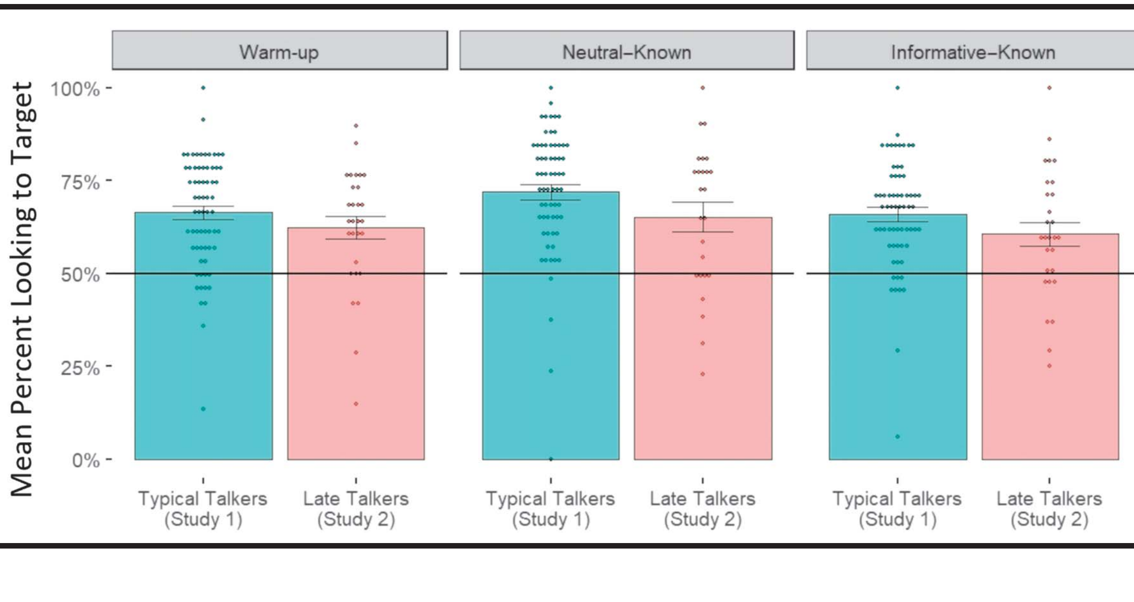
We then assessed late talkers' verb-based prediction beginning with the Informative–Known trials. Late talkers showed significantly greater looking to the target than to the distracter following the onset of the verb, $M = 0.61$, $SD = 0.17$, $t(27) = 3.05$, $p = .005$. Typical talkers looked somewhat more to the target on Informative–Known trials than late talkers did (see Figure 5); this effect approached significance, $t(86) = 1.85$, $p = .068$. As in Study 1, we also evaluated Informative–Known trials separately for "eat" and "wear" sentences. Like typical talkers, late talkers showed above-chance looking to the target on "eat" trials, $M = 0.69$, $SD = 0.30$, $t(25) = 3.03$, $p = .006$, but not on "wear" trials, $M = 0.49$, $SD = 0.23$, $t(24) = 0.13$, $p = .90$. The effect of verb was significant, $t(70) = 6.86$, $p < .0001$.

To confirm that late talkers were using the verb to identify the target referent, we compared their looking to the target on Informative–Known and Neutral–Known trials during the same window (233–2,000 ms after verb onset [post-verb]). This analysis revealed that late talkers looked significantly more to the target referent during the post-verb window when the verb was informative (e.g., "eat") than when it was neutral (e.g., "see"), $t(22) = 4.14$, $p = .0004$ (see Figure 3, right panel). Thus, like typical talkers in Study 1, late talkers successfully used an informative verb to rapidly infer the referent of the upcoming noun.

Word Learning

Next, we examined Informative–Novel trials, which asked children to use verbs to learn new nouns (see Figure 4). Overall, late talkers tended to look to the target at numerically above-chance rates, $M = 0.56$, $SD = 0.19$; this approached significance, $t(26) = 1.77$, $p = .089$. We again

Figure 5. Proportion of time looking to the target across word recognition trials. Both typical and late talkers showed above-chance looking to the target on trials assessing recognition of known nouns (Warm-up and Neutral–Known trial types) and known verbs (Informative–Known trial type), ps < .05. Typical talkers looked significantly more to the target than did the late talkers only on Informative–Known trials, $p < .05$, indicating more accurate verb-based prediction. Typical-talker results presented here collapse across Study 1's task versions because no differences emerged between Massed and Interleaved tasks, ps > .1.



observed an effect of verb, $t(27) = 2.17$, $p = .04$, with late talkers showing a stronger, albeit not significant, preference for the target on "eat" trials, $M = 0.60$, $SD = 0.29$, $t(25) = 1.62$, $p = .12$, than on "wear" trials, $M = 0.46$, $SD = 0.26$, $t(19) = 0.63$, $p = .53$. Comparing performance across studies, while controlling for the effect of verb, revealed that typical talkers looked significantly more to the target than late talkers did, $t(77) = 2.69$, $p = .009$ (see Figure 4, left panel, blue and red bars). Thus, late-talking children were less successful than typical talkers in using a semantically informative verb to identify the referent of the novel noun.

Next, we examined Neutral–Novel trials (e.g., "Find the dax") to assess whether late talkers learned the meanings of the novel nouns introduced in the Informative–Novel trials. Late talkers showed overall greater looking to the target referent than to the distracter, $M = 0.63$, $SD = 0.22$, $t(24) = 2.78$, $p = .01$. More specifically, children showed above-chance looking to the target when the noun had been introduced with "eat," $M = 0.66$, $SD = 0.24$, $t(23) = 3.24$, $p = .004$, but not when it had been introduced with "wear," $M = 0.58$, $SD = 0.30$, $t(17) = 1.12$, $p = .28$. After controlling for verb, no significant difference emerged between late and typical talkers on Neutral–Novel trials,[2] $t(38) = 1.46$, $p = .24$ (see Figure 4, right panel, blue and red bars). Children in both groups showed limited success on this component of the task: learning novel nouns but only when using the informative verb "eat."

Relations Between Novel Word Exposure and Test Trials

Next, we considered the relation between children's performance on trials teaching novel words (i.e., Informative–Novel trials) and trials testing children's understanding of these novel words (i.e., Neutral–Novel trials). Although we have thus far analyzed these two trial types separately, they are necessarily interrelated. Children are asked to infer the referent of a novel noun on Informative–Novel trials, and they are then tested on this novel noun–referent mapping on the subsequent Neutral–Novel trials. This presents an opportunity to assess the relation between verb-based prediction and word learning.

First, we constructed a linear model of children's average looking time to the target on Neutral–Novel trials in the Massed version of the task, predicted by group and average performance on learning trials (i.e., looking time to the target on Informative–Novel trials). This analysis revealed no significant effects of group, $t(55) = 0.76$, $p = .45$, or performance on Informative–Novel trials, $t(55) = 1.22$, $p = .23$. However, the interaction between the two factors was significant,[3] $t(55) = 2.03$, $p = .047$. This interaction reflects a stronger relation between performance on Informative–Novel and Neutral–Novel trials in typical talkers, as compared with that in late talkers (see Figure 6). Strikingly, performance on Informative–Novel and Neutral–Novel trials was essentially unrelated for late talkers, $r(23) = -.09$, $p = .67$, but positively associated for typical talkers, $r(32) = .49$, $p = .003$.

Finally, we conducted two additional analyses to test whether this association represented a genuine relation between exposure and test. First, we asked whether the learning–test correlation could be attributed to a general effect of children's engagement with the task. If this were the case, then high performance on any trial type should correlate with high performance on any other. However, this was not the case: Children's performance on Warm-up trials did not predict their performance on Neutral–Novel trials, $t(57) = 0.84$, $p = .41$, nor was there a Group × Warm-up performance interaction, $t(57) = 0.13$, $p = .90$. Thus, the learning–test correlation cannot be attributed to individual differences in engagement and general task performance.

Next, we asked whether the learning–test correlation could reflect individual children's biases for a particular image. If so, a child's preference during the *baseline* period of Informative–Novel trials, before the sentence begins, should also predict their looking preference on Neutral–Novel trials. However, baseline preferences did not significantly predict looking on Neutral–Novel trials, $t(56) = 0.12$, $p = .91$, nor did this interact with group, $t(56) = 0.92$, $p = .36$. This suggests a unique relationship between performance on the Informative–Novel ("teaching") and Neutral–Novel ("test") trials: Better online verb-based referent prediction appears to lead to better novel noun learning. However, this relationship is only present in typical talkers. We return to this point in the discussion.

Processing Efficiency

As in Study 1, we analyzed processing efficiency separately for trials where the disambiguating word was either a noun or a verb (see Figure 7). Like typical talkers, late talkers were somewhat slower to shift to the target on verb trials than on noun trials, but this difference did not reach significance, $B = 159$, $SE = 106$,

[3]At the suggestion of a reviewer, we also ran a version of this analysis including only those participants who contributed three or more Informative–Novel and Neutral–Novel trials, yielding more precise estimates of individual children's performance. The results of this analysis were largely similar, although the interaction did not reach significance, $t(26) = 1.84$, $p = .077$, likely due to the smaller sample size.

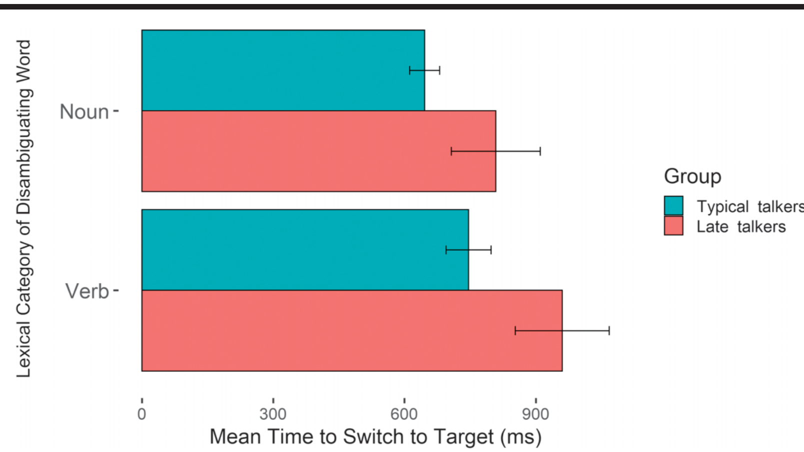
Figure 6. Relationship between looking on Informative–Novel ("teaching") and Neutral–Novel ("test") trials. Typical talkers showed a stronger, more positive relationship between looking on Informative–Novel ("You can eat the dax") and Neutral–Novel ("Find the dax") trials than their late-talker peers, $p = .047$.

$t(19) = 1.50$, $p = .15$. Late talkers were also less efficient than typical talkers in recognizing familiar nouns and using verbs to predict upcoming referents. For nouns (i.e., Warm-up and Neutral–Known trials), late talkers shifted toward the target more slowly, $M = 809$ ms, $SD = 519$, than did the typical talkers, $B = 148$, $SE = 68$, $t(61) = 2.18$, $p = .033$. A similar pattern emerged on verb trials (i.e., Informative–Known and Informative–Novel). Late talkers were significantly slower to shift to the target image, $M = 935$ ms, $SD = 581$, than typical talkers, $B = 183$, $SE = 81$, $t(47) = 2.26$, $p = .028$, controlling for verb. In sum, late-talking children used both nouns and verbs to identify the target referent but were slower to do so than typical talkers.

General Discussion

We conducted a comprehensive assessment of online lexical processing abilities in 2-year-old typical talkers (Study 1) and late talkers (Study 2), testing their comprehension of familiar nouns and verbs as well as their ability to use those familiar verbs to learn and retain novel

Figure 7. Mean switch time across lexical categories and language groups. Switch times indicate the speed with which children switched their gaze from the distracter referent to the target referent after the onset of the cueing word (verb or noun). Typical talkers shifted more quickly in response to nouns than to verbs, $p = .030$. For both nouns and verbs, late talkers were slower to shift to the target than typical talkers, $ps < .05$.



noun–object mappings. Five key findings emerged: (a) Late and typical talkers showed accurate and rapid recognition of familiar nouns. (b) Similarly, both late and typical talkers used the semantic information in a familiar verb to anticipate an upcoming noun. This was evidenced by directing their visual attention to the target more rapidly when hearing sentences with an informative verb ("You can *eat* the apple") than when hearing sentences with a neutral verb ("You can *see* the apple"). (c) Late and typical talkers showed similar levels of success in using semantically informative verbs to infer the referent of a *novel* noun, although for both groups, this effect was observed only with the verb "eat" and not with "wear." (d) Late talkers were less efficient than typical talkers in recognizing familiar nouns and in using familiar verbs to anticipate an upcoming noun, as evidenced by their longer latencies to shift to the referent on both noun and verb trials. (e) Among typical talkers, greater looking to the target on novel noun learning trials ("You can eat the dax") predicted accuracy on test trials ("Find the dax"), suggesting a relationship between verb-based prediction and word learning; this relationship was not observed in late talkers. Indeed, despite late talkers' less efficient processing of the semantically informative verbs, they were equally successful in learning the novel noun–referent pairings.

Overall, then, the results reported here lend additional support to the link between vocabulary knowledge and lexical processing efficiency in toddlers, indicating that late talkers do not simply know fewer words but also process the words they know more slowly in online speech (Borovsky et al., 2012; Fernald et al., 2006; Mani & Huettig, 2012). Late talkers, who have atypically small productive vocabularies, did not differ from typical talkers in their accuracy in recognizing familiar nouns (all of which were produced by a majority of 24-month-olds), but they were significantly slower to orient to the nouns' referents (see also Fernald & Marchman, 2012). In addition, late talkers were both slower and less accurate than typical talkers in using a semantically informative verb to identify the referent of an upcoming noun, whether the noun was novel or familiar. This suggests that at 2 years of age, late talkers may differ more from typical talkers in their verb-based prediction than in their processing of familiar nouns. Thus, although early assessments of language ability in the first 2 years often focus on nouns, more complex linguistic tasks such as verb-based prediction may provide a better diagnostic tool. Indeed, evidence suggests that the association between vocabulary and noun processing efficiency weakens throughout the third year of life (Koenig et al., 2020; Mahr & Edwards, 2018; Peter et al., 2019). In contrast, toddlers' processing of verbs, which is linked to later grammatical development, may remain a crucial predictor of language outcomes during this period (cf. Hadley et al., 2016). Recent evidence from van Alphen et al. (2021) is consistent with this possibility, suggesting that 3- to 4-year-old children suspected of having DLD are slower to process verbs, but not nouns, compared to their typically developing peers.

This raises a compelling question about the mechanisms underlying online verb processing and verb-based prediction more generally. Children's successful use of verbs here has at least two possible interpretations. First, children may have recruited the verbs' selectional restrictions to predict the upcoming noun. That is, they might have recognized that the verb "eat" requires an edible item as its object (Chomsky, 1965; Grimshaw, 1979; Resnik, 1996) or at least that the object of "eat" is very likely to be edible (Warren et al., 2015). Alternatively, children's performance may have stemmed from a coarser strategy, such as semantic association. Merely hearing "eat" might prime children to attend more to semantically related objects such as food, as well as other items such as mouths or forks (cf. Bergelson & Aslin, 2017). In future work, it will be important to assess whether the effects observed here reflect differences in children's ability to recruit verbs' selectional restrictions to predict an upcoming noun (Mani & Huettig, 2012) and/or differences in the strength of their semantic associations or the structure of their lexico-semantic networks (Borovsky et al., 2016b).

In either case, it is clear that an ability to use knowledge of a verb's meaning to anticipate the referent of an upcoming noun is a critical component of language processing. In this task, children successfully used verbs not only to identify familiar objects but also to infer the meanings of novel words, suggesting that this ability is sufficiently robust and flexible to support word learning (Goodman et al., 1998). However, both typical and late talkers were only successful in learning the novel noun referring to the food item (presented as the object of "eat") and not the clothing item (presented as the object of "wear"), a difference that was unanticipated (cf. Borovsky et al., 2016b). This difference could stem from multiple sources. First, perhaps our clothing objects, particularly the "novel" object we chose (a kimono), did not present as obviously wearable to our children. Alternatively, perhaps these findings indicate a disparity in children's use of these verbs. In particular, the verb "eat" may have presented a particularly accessible set of selectional restrictions: MacRoy-Higgins et al. (2016) compared 2-year-old late and typical talkers' vocabularies across 22 different semantic categories derived from the MCDI and found that food words were one of the most consistently well-known categories across late talkers, age-matched typical talkers, and a vocabulary-matched control group, as well as the most well-represented semantic category of nouns in late talkers' vocabularies. In contrast, the verb "wear" could pose multiple difficulties for toddlers. Although some work has shown successful use of its selectional restrictions (Borovsky et al., 2016b), the category of clothing is less well represented in 2-year-olds' productive vocabularies for both late and typical talkers (Borovsky et al., 2016a; MacRoy-Higgins et al., 2016). In addition, although "wear" and "eat" are similarly frequent in child-directed speech at 24 months of age, "wear" becomes less frequent over the course of the third year and is far less frequent in children's own speech (MacWhinney, 2014; Sanchez et al., 2019). In part, this may be because "wear" is synonymous with another frequent verb phrase, namely, "put on," which may restrict its use to certain contexts (e.g., children might hear "*Wear* your glasses," but "*Put on* your pants"). "Wear" is also homophonous with another early-acquired word, namely, "where," which might interfere with children's processing. Finally, because children were learning novel words for two novel objects concurrently, it is possible that if children successfully learned the food referent's label first, they simply attended more to that object knowing it had been previously labeled and discussed (cf. Twomey & Westermann, 2018). This tendency, or perhaps a more general food saliency bias, may have masked a slower but potentially still successful learning process for the clothing item's label. In sum, although we provide evidence that typical and late talkers perform similarly in processing "eat" and "wear," which are well-known and less well-known verbs, respectively, perhaps differences would emerge in a middle ground, particularly with verbs for which late talkers tend to know fewer of their objects. In future work, we hope to assess children's performance with a greater variety of verbs.

Although these findings reveal a number of differences in online lexical processing between typical and late talkers, they also show notable similarities. Most critically, typical and late talkers were similarly successful in learning the novel words. Previous studies had examined late talkers' novel word learning primarily in fast-mapping procedures (Ellis Weismer, 2007; MacRoy-Higgins & Montemarano, 2016; though see Ellis et al., 2015), which feature unambiguous referents. Instead, word learning in this study occurred in a context in which the novel words were not introduced ostensively but rather required toddlers to use a familiar verb to identify the novel noun's referent from an otherwise ambiguous display. Late talkers' success suggests that their slower processing efficiency (observed here on Informative–Known and Informative–Neutral trials) need not necessarily impede their learning and retention of the word–referent mapping, at least when the interval between learning and test is short. This is reflected in the differential relation between "teaching" and "test" trials for typical and late talkers: While toddlers' performance on Informative–Novel and Neutral–Novel trials was correlated in typical talkers, this was not true in late talkers. Late talkers may have initially struggled to make the correct inference on learning trials, leading to less systematic looking on these trials, yet they ultimately used the verb to identify the correct referent and learned the noun–object mapping.

This finding raises intriguing questions about the link between online word recognition and word learning. On one hand, the results from typical talkers demonstrate a specific link between word learning and retention, suggesting that children who are more efficient at using a semantically related verb to identify the referent of an upcoming noun are also more successful in learning and retaining the mapping between that (novel) noun and its referent. On the other hand, this relationship was not observed among late talkers, who, despite being slower to predict the target during learning, appeared to learn the novel noun–object mapping as well as typical talkers. Strikingly, then, late talkers' delay in online processing did not appear to translate into less robust learning. One possibility is that the underlying learning–test relationship is similar for late and typical talkers, but late talkers' shift to the target simply occurred later. If that were the case, late talkers' looking during the naming window would be a less accurate measure of their eventual learning than for typical talkers. Another possibility is that the word learning test here was insufficiently sensitive to detect differences between typical and late talkers. Previous work found impairments in word learning in children with language disorders that emerged only when the task demands were higher (e.g., when there were a greater number of novel words and/or when the phonological forms of different novel words are more similar; Gordon et al., 2021; MacRoy-Higgins et al., 2013; Weismer & Hesketh, 1996). Thus, although differences in word learning were not detected in this task, they might emerge in other learning environments. Finally, as noted above, there are multiple ways for children to arrive at the correct word–referent mapping: Perhaps late talkers used different strategies than the typical talkers for succeeding in this task, leading to a different or more variable pattern of eye movements (cf. McGregor et al., 2022). Further research is needed to tease apart these possible explanations. However, our results suggest that we should be cautious in characterizing the potential consequences of delays in online language processing for language learning.

These findings have important implications for early identification of children with language disorders. Although the current analyses focused on group differences between late and typical talkers, future work might explore individual differences in both groups in the relation between language processing and later outcomes, including subsequent vocabulary growth and grammatical competence. For instance, although late talkers' language outcomes are quite heterogeneous as a group, it is possible that those late talkers who struggle with online language comprehension are also substantially more likely to experience persistent language impairments (cf. Ellis & Thal, 2008; Fernald & Marchman, 2012). More broadly, online language processing might play an important role in the early identification of a variety of future language impairments, such as those associated with autism spectrum disorder (cf. Prescott et al., 2022; Venker et al., 2019; Zhou et al., 2019). By gathering robust measures of children's online processing of both nouns and verbs, we may be able to make better and earlier predictions about children's future language learning trajectories. A major goal of the broader longitudinal study of which this study is a part is to examine how a variety of measures, including neural processing of language, cognitive ability, home language environment, and child irritability, can best reveal which late talkers are likely to catch up with peers and which will continue to struggle and would benefit from early intervention. This study suggests that language processing assessments offer a promising way of detecting language impairments early in life, particularly if these assessments include not only noun recognition but also more complex skills such as verb-based prediction. These results also highlight late talkers' success in using semantic information to learn new words, despite initial processing challenges, indicating that researchers and clinicians may be able to leverage these word learning skills in creating effective interventions.

Data Availability Statement

The data sets generated and analyzed during this study are available from the corresponding author upon request and after additional analyses are completed.

Ethics Approval Statement

All research was approved by the Northwestern University Institutional Review Board, and parents gave informed consent for their children's participation.

Acknowledgments

Research reported in this article was supported by National Institute on Deafness and Other Communication Disorders Grants R01DC016273 (principal investigators [PIs]: Elizabeth S. Norton and Lauren S. Wakschlag) and R21DC018357 (PI: Adriana Weisleder), National Institute of Mental Health Grant R01MH107652 (PI: Lauren S. Wakschlag), and Eunice Kennedy Shriver National Institute of Child Health and Human Development Grants F32HD103448 (Alexander LaTourrette) and R01HD083310 (PI: Sandra Waxman). The authors would like to thank Amy Biel, Kiera Cook, Kaitie Fredian, Ewa Gut, Alex Harpole, Emily Harriott, Sheila Krogh-Jespersen, Jessica Page, Amelie Petitclerc, Debby Zemlock, Yudong Zhang, and Anne Zola for their help in collecting and organizing data, as well as Emma Smyles, Courtney Goldberg, and Lauren Miller for their help in creating stimuli. Special thanks also go to Pamela Hadley for insightful comments on earlier drafts.

References

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

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

Andreu, L., Sanz-Torrent, M., & Guàrdia-Olmos, J. (2012). Auditory word recognition of nouns and verbs in children with specific language impairment (SLI). *Journal of Communication Disorders, 45*(1), 20–34. https://doi.org/10.1016/j.jcomdis.2011.09.003

Andreu, L., Sanz-Torrent, M., Legaz, L. B., & MacWhinney, B. (2012). Effect of verb argument structure on picture naming in children with and without specific language impairment (SLI). *International Journal of Language & Communication Disorders, 47*(6), 637–653. https://doi.org/10.1111/j.1460-6984.2012.00170.x

Arias-Trejo, N., & Plunkett, K. (2013). What's in a link: Associative and taxonomic priming effects in the infant lexicon. *Cognition, 128*(2), 214–227. https://doi.org/10.1016/j.cognition.2013.03.008

Babineau, M., de Carvalho, A., Trueswell, J., & Christophe, A. (2021). Familiar words can serve as a semantic seed for syntactic bootstrapping. *Developmental Science, 24*(1), Article e13010. https://doi.org/10.1111/desc.13010

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

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

Benitez, V. L., Zettersten, M., & Wojcik, E. (2020). The temporal structure of naming events differentially affects children's and adults' cross-situational word learning. *Journal of Experimental Child Psychology, 200,* 104961. https://doi.org/10.1016/j.jecp.2020.104961

Bergelson, E., & Aslin, R. (2017). Semantic specificity in one-year-olds' word comprehension. *Language Learning and Development, 13*(4), 481–501. https://doi.org/10.1080/15475441.2017.1324308

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016a). Lexical leverage: Category knowledge boosts real-time novel word recognition in 2-year-olds. *Developmental Science, 19*(6), 918–932. https://doi.org/10.1111/desc.12343

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016b). Semantic structure in vocabulary knowledge interacts with lexical and sentence processing in infancy. *Child Development, 87*(6), 1893–1908. https://doi.org/10.1111/cdev.12554

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology, 112*(4), 417–436. https://doi.org/10.1016/j.jecp.2012.01.005

Chomsky, N. (1965). *Aspects of the theory of syntax* (Vol. 11). MIT Press.

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

Dale, P. S., Price, T. S., Bishop, D. V. M., & Plomin, R. (2003). Outcomes of early language delay: I. Predicting persistent and transient language difficulties at 3 and 4 years. *Journal of Speech, Language, and Hearing Research, 46*(3), 544–560. https://doi.org/10.1044/1092-4388(2003/044)

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, E. M., & Thal, D. J. (2008). Early language delay and risk for language impairment. *SIG 1 Perspectives on Language Learning and Education, 15*(3), 93–100. https://doi.org/10.1044/lle15.3.93

Ellis Weismer, S. (2007). Typical talkers, late talkers, and children with specific language impairment: A language endowment spectrum? In R. Paul (Ed.), *Language disorders from a developmental perspective: Essays in honor of Robin S. Chapman* (pp. 83–101). Psychology Press.

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

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2006). *MacArthur–Bates Communicative Development Inventories: User's guide and technical manual*. Brookes.

Ferguson, B., Graf, E., & Waxman, S. R. (2014). Infants use known verbs to learn novel nouns: Evidence from 15- and 19-month-olds. *Cognition, 131*(1), 139–146. https://doi.org/10.1016/j.cognition.2013.12.014

Ferguson, B., Graf, E., & Waxman, S. R. (2018). When *veps* cry: Two-year-olds efficiently learn novel words from linguistic contexts alone. *Language Learning and Development, 14*(1), 1–12. https://doi.org/10.1080/15475441.2017.1311260

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

Fernald, A., Zangl, R., Portillo, A. L., & Marchman, V. A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds.), *Developmental psycholinguistics: On-line methods in children's language processing* (Vol. 44, p. 97). John Benjamins. https://doi.org/10.1075/lald.44.06fer

Fisher, C., Gertner, Y., Scott, R. M., & Yuan, S. (2010). Syntactic bootstrapping. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(2), 143–149. https://doi.org/10.1002/wcs.17

Fisher, C., Jin, K., & Scott, R. M. (2020). The developmental origins of syntactic bootstrapping. *Topics in Cognitive Science, 12*(1), 48–77. https://doi.org/10.1111/tops.12447

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

Frank, M. C., Tenenbaum, J. B., & Fernald, A. (2013). Social and discourse contributions to the determination of reference in cross-situational word learning. *Language Learning and Development, 9*(1), 1–24. https://doi.org/10.1080/15475441.2012.707101

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition, 73*(2), 135–176. https://doi.org/10.1016/S0010-0277(99)00036-0

Goodman, J. C., McDonough, L., & Brown, N. B. (1998). The role of semantic context and memory in the acquisition of novel nouns. *Child Development, 69*(5), 1330–1344. https://doi.org/10.2307/1132269

Gordon, K. R., Storkel, H. L., Lowry, S. L., & Ohlmann, N. B. (2021). Word learning by preschool-age children with developmental language disorder: Impaired encoding and robust consolidation during slow mapping. *Journal of Speech, Language, and Hearing Research, 64*(11), 4250–4270. https://doi.org/10.1044/2021_JSLHR-21-00046

Graf Estes, K., Edwards, J., & Saffran, J. R. (2011). Phonotactic constraints on infant word learning. *Infancy, 16*(2), 180–197. https://doi.org/10.1111/j.1532-7078.2010.00046.x

Grimshaw, J. (1979). Complement selection and the lexicon. *Linguistic Inquiry, 10*(2), 279–326.

Hadley, P. A. (2006). Assessing the emergence of grammar in toddlers at risk for specific language impairment. *Seminars in Speech and Language, 27*(3), 173–186. https://doi.org/10.1055/s-2006-948228

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

Henrichs, J., Rescorla, L., Schenk, J. J., Schmidt, H. G., Jaddoe, V. W., Hofman, A., Raat, H., Verhulst, F. C., & Tiemeier, H. (2011). Examining continuity of early expressive vocabulary development: The Generation R Study. *Journal of Speech, Language, and Hearing Research, 54*(3), 854–869. https://doi.org/10.1044/1092-4388(2010/09-0255)

Horvath, S., & Arunachalam, S. (2021). Repetition versus variability in verb learning: Sometimes less is more. *Journal of Speech, Language, and Hearing Research, 64*(11), 4235–4249. https://doi.org/10.1044/2021_JSLHR-21-00091

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

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

Koenig, A., Arunachalam, S., & Saudino, K. J. (2020). Lexical processing of nouns and verbs at 36 months of age predicts concurrent and later vocabulary and school readiness. *Journal of Cognition and Development, 21*(5), 670–689. https://doi.org/10.1080/15248372.2020.1802277

Krok, W., Norton, E. S., Buchheit, M. K., Harriott, E. M., Wakschlag, L., & Hadley, P. A. (2022). Using animated action scenes to remotely assess sentence diversity in toddlers. *Topics in Language Disorders, 42*(2), 156–172. https://doi.org/10.1097/TLD.0000000000000280

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software, 82*(13), 1–26.

Lany, J. (2018). Lexical-processing efficiency leverages novel word learning in infants and toddlers. *Developmental Science, 21*(3), Article e12569. https://doi.org/10.1111/desc.12569

Luke, S. G. (2017). Evaluating significance in linear mixed-effects models in R. *Behavior Research Methods, 49*(4), 1494–1502. https://doi.org/10.3758/s13428-016-0809-y

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., Schwartz, R. G., Shafer, V. L., & Marton, K. (2013). Influence of phonotactic probability/neighbourhood density on lexical learning in late talkers. *International Journal of Language & Communication Disorders, 48*(2), 188–199. https://doi.org/10.1111/j.1460-6984.2012.00198.x

MacRoy-Higgins, M., Shafer, V. L., Fahey, K. J., & Kaden, E. R. (2016). Vocabulary of toddlers who are late talkers. *Journal of Early Intervention, 38*(2), 118–129. https://doi.org/10.1177/1053815116637620

MacWhinney, B. (2014). *The CHILDES Project: Tools for analyzing talk, Volume II: The database*. Psychology Press.

Mahr, T., & Edwards, J. (2018). Using language input and lexical processing to predict vocabulary size. *Developmental Science, 21*(6), Article e12685. https://doi.org/10.1111/desc.12685

Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4), 843–847. https://doi.org/10.1037/a0029284

McGregor, K. K., Smolak, E., Jones, M., Oleson, J., Eden, N., Arbisi-Kelm, T., & Pomper, R. (2022). What children with developmental language disorder teach us about cross-situational word learning. *Cognitive Science, 46*(2), Article e13094. https://doi.org/10.1111/cogs.13094

Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences, 108*(22), 9014–9019. https://doi.org/10.1073/pnas.1105040108

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

Mullen, E. M. (1995). *Mullen Scales of Early Learning*. AGS.

Norton, E. S., Manning, B. L., Harriott, E. M., Nikolaeva, J. I., Nyabingi, O. S., Fredian, K. M., Page, J. M., McWeeny, S., Krogh-Jespersen, S., MacNeill, L. A., Roberts, M. Y., & Wakschlag, L. S. (2022). Social EEG: A novel neurodevelopmental approach to studying brain–behavior links and brain-to-brain synchrony during naturalistic toddler–parent interactions. *Developmental Psychobiology, 64*(3), Article e22240. https://doi.org/10.1002/dev.22240

Paul, R., Murray, C., Clancy, K., & Andrews, D. (1997). Reading and metaphonological outcomes in late talkers. *Journal of Speech, Language, and Hearing Research, 40*(5), 1037–1047. https://doi.org/10.1044/jslhr.4005.1037

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. https://doi.org/10.1044/2019_JSLHR-L-ASTM-18-0234

Peter, M. S., Durrant, S., Jessop, A., Bidgood, A., Pine, J. M., & Rowland, C. F. (2019). Does speed of processing or vocabulary size predict later language growth in toddlers? *Cognitive Psychology, 115,* 101238. https://doi.org/10.1016/j.cogpsych.2019.101238

Peters, R., & Borovsky, A. (2019). Modeling early lexico-semantic network development: Perceptual features matter most. *Journal of Experimental Psychology: General, 148*(4), 763–782. https://doi.org/10.1037/xge0000596

Prescott, K. E., Mathée-Scott, J., Reuter, T., Edwards, J., Saffran, J., & Ellis Weismer, S. (2022). Predictive language processing in young autistic children. *Autism Research, 15*(5), 892–903. https://doi.org/10.1002/aur.2684

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

Reilly, S., Wake, M., Ukoumunne, O. C., Bavin, E., Prior, M., Cini, E., Conway, L., Eadie, P., & Bretherton, L. (2010). Predicting language outcomes at 4 years of age: Findings from Early Language in Victoria Study. *Pediatrics, 126*(6), e1530–e1537. https://doi.org/10.1542/peds.2010-0254

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

Rescorla, L. (2002). Language and reading outcomes to age 9 in late-talking toddlers. *Journal of Speech, Language, and Hearing Research, 45*(2), 360–371. https://doi.org/10.1044/1092-4388(2002/028)

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., Dahlsgaard, K., & Roberts, J. (2000). Late-talking toddlers: MLU and IPSyn outcomes at 3;0 and 4;0. *Journal of Child Language, 27*(3), 643–664. https://doi.org/10.1017/S0305000900004232

Rescorla, L., & Dale, P. S. (2013). *Late talkers: Language development, interventions, and outcomes* (1st ed.). Brookes.

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

Resnik, P. (1996). Selectional constraints: An information-theoretic model and its computational realization. *Cognition, 61*(1–2), 127–159. https://doi.org/10.1016/S0010-0277(96)00722-6

Rice, M. L., Taylor, C. L., & Zubrick, S. R. (2008). Language outcomes of 7-year-old children with or without a history of late language emergence at 24 months. *Journal of Speech, Language, and Hearing Research, 51*(2), 394–407. https://doi.org/10.1044/1092-4388(2008/029)

Sanchez, A., Meylan, S. C., Braginsky, M., MacDonald, K. E., Yurovsky, D., & Frank, M. C. (2019). childes-db: A flexible and reproducible interface to the Child Language Data Exchange System. *Behavior Research Methods, 51*(4), 1928–1941. https://doi.org/10.3758/s13428-018-1176-7

Schwab, J. F., & Lew-Williams, C. (2016). Repetition across successive sentences facilitates young children's word learning. *Developmental Psychology, 52*(6), 879–886. https://doi.org/10.1037/dev0000125

Schwab, J. F., & Lew-Williams, C. (2020). Discontinuity of reference hinders children's learning of new words. *Child Development, 91*(1), e29–e41. https://doi.org/10.1111/cdev.13189

Smith, L., & 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

Swingley, D., Pinto, J. P., & Fernald, A. (1999). Continuous processing in word recognition at 24 months. *Cognition, 71*(2), 73–108. https://doi.org/10.1016/S0010-0277(99)00021-9

Twomey, K. E., & Westermann, G. (2018). Learned labels shape pre-speech infants' object representations. *Infancy, 23*(1), 61–73. https://doi.org/10.1111/infa.12201

Valleau, M. J., Konishi, H., Golinkoff, R. M., Hirsh-Pasek, K., & Arunachalam, S. (2018). An eye-tracking study of receptive verb knowledge in toddlers. *Journal of Speech, Language, and Hearing Research, 61*(12), 2917–2933. https://doi.org/10.1044/2018_JSLHR-L-17-0363

van Alphen, P., Brouwer, S., Davids, N., Dijkstra, E., & Fikkert, P. (2021). Word recognition and word prediction in preschoolers with (a suspicion of) a developmental language disorder: Evidence from eye tracking. *Journal of Speech, Language, and Hearing Research, 64*(6), 2005–2021. https://doi.org/10.1044/2021_JSLHR-20-00227

Venker, C. E., Edwards, J., Saffran, J. R., & Ellis Weismer, S. (2019). Thinking ahead: Incremental language processing is associated with receptive language abilities in preschoolers with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 49*(3), 1011–1023. https://doi.org/10.1007/s10803-018-3778-4

Verhoeven, L., & van Balkom, H. (2003). *Classification of developmental language disorders: Theoretical issues and clinical implications*. Psychology Press. https://doi.org/10.4324/9781410609021

Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children's cross-situational word learning, memory, and language abilities. *Journal of Memory and Language, 93,* 217–230. https://doi.org/10.1016/j.jml.2016.10.001

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition, 127*(3), 375–382. https://doi.org/10.1016/j.cognition.2013.02.015

Wakschlag, L. S., Roberts, M. Y., Flynn, R. M., Smith, J. D., Krogh-Jespersen, S., Kaat, A. J., Gray, L., Walkup, J., Marino, B. S., & Norton, E. S. (2019). Future directions for early childhood prevention of mental disorders: A road map to mental health, earlier. *Journal of Clinical Child & Adolescent Psychology, 48*(3), 539–554. https://doi.org/10.1080/15374416.2018.1561296

Warren, T., Milburn, E., Patson, N. D., & Dickey, M. W. (2015). Comprehending the impossible: What role do selectional restriction violations play? *Language, Cognition and Neuroscience, 30*(8), 932–939. https://doi.org/10.1080/23273798.2015.1047458

Weismer, S. E., & Hesketh, L. J. (1996). Lexical learning by children with specific language impairment: Effects of linguistic input presented at varying speaking rates. *Journal of Speech and Hearing Research, 39*(1), 177–190. https://doi.org/10.1044/jshr.3901.177

Westerlund, M., Berglund, E., & Eriksson, M. (2006). Can severely language delayed 3-year-olds be identified at 18 months? Evaluation of a screening version of the MacArthur–Bates Communicative Development Inventories. *Journal of Speech, Language, and Hearing Research, 49*(2), 237–247. https://doi.org/10.1044/1092-4388(2006/020)

Zhou, P., Zhan, L., & Ma, H. (2019). Predictive language processing in preschool children with autism spectrum disorder: An eye-tracking study. *Journal of Psycholinguistic Research, 48*(2), 431–452. https://doi.org/10.1007/s10936-018-9612-5

Zubrick, S. R., Taylor, C. L., Rice, M. L., & Slegers, D. W. (2007). Late language emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates. *Journal of Speech, Language, and Hearing Research, 50*(6), 1562–1592. https://doi.org/10.1044/1092-4388(2007/106)