Repetition vs. variability in verb learning: Sometimes less is more

Sabrina Horvath[1] & Sudha Arunachalam[2]

[1]Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, IN

[2]Department of Communicative Sciences and Disorders, New York University, New York, NY

Corresponding Author: Sabrina Horvath

horvath7@purdue.edu

Lyles-Porter Hall, 715 Clinic Drive, West Lafayette, IN 47907

Conflict of Interest: The authors have no conflicts of interest to disclose

Funding Sources: This research was supported by NSF BCS 1748826, NIH R01DC016592, and an ASHFoundation New Century Scholars Research Grant to SA; an ASHFoundation Student Research Grant in Early Childhood Language Development (supported by the Arlene M. and Noel D. Matkin Memorial Fund) to SH, and a Boston University Dudley Allen Sargent Research Fund grant to SH.

Abstract

Purpose: This study examined whether 2-year-olds are better able to acquire novel verb meanings when they appear in varying linguistic contexts, including both content nouns and pronouns, as compared to when the contexts are consistent, including only content nouns. Additionally, differences between typically developing toddlers and late talkers were explored.

Method: Forty-seven English-acquiring 2-year-olds (*n* = 14 late talkers, *n* = 33 typically developing) saw scenes of actors manipulating objects. These actions were labeled with novel verbs. In the Varied condition, children heard sentences containing both content nouns and pronouns (e.g., "The girl is ziffing the truck. She is ziffing it!"). In the Consistent condition, children heard the verb an equal number of times, but only with content nouns (e.g., "The girl is ziffing the truck. The girl is ziffing the truck!"). At test, children were shown two new scenes and were asked to find the novel verb's referent. Children's eye gaze was analyzed as a measure of learning.

Results: Mixed-effects regression analyses revealed that children looked more toward the correct scene in the Consistent condition than the Varied condition. This difference was more pronounced for late talkers than for typically developing children.

Conclusions: To acquire an initial representation of a new verb's meaning, children, particularly late talkers, benefit more from hearing the verb in consistent linguistic contexts than in varying contexts.

WC: 222

Keywords: Verbs, Syntactic Bootstrapping, Variability, Late Talkers, Word Learning

Introduction

Learning the meanings of words is one of the most impressive achievements of early childhood. It requires children to integrate social, cognitive, and linguistic skills to mine the environment for cues to a new word's meaning. All of these skills are important because different kinds of words may be learned by different word learning mechanisms. Children's early vocabularies are typically dominated by nouns that refer to concrete entities such as "shoe" or "cookie" (e.g., Gentner, 1982; Gentner & Boroditsky, 2001; Longobardi et al., 2017; Nelson, 1973). A good learning situation for such vocabulary items may involve the caregiver sharing attention with the child while pointing to the object denoted by the noun (e.g., Tomasello & Farrar, 1986). However, this type of learning situation is less likely to be helpful for learning verbs, which typically denote events that involve one or more participants engaged in some action or some relation to each other (e.g., Gentner, 1978). Learning verbs (and nouns that denote events; Arunachalam & He, 2018) is thought to be a particular challenge because children may need to rely on the linguistic context in which the word appears to identify which aspect of the event is being labeled (e.g., Gleitman, 1990; Landau & Gleitman, 1985; Naigles, 1990). In a classic demonstration, Fisher et al. (1994) found that, while viewing events involving two participants, 3- and 4-year-old children used a novel verb's linguistic context to identify its meaning. For example, given a scene depicting a monkey on a rabbit's shoulders, children who heard "The monkey is gorping the rabbit" were more likely to think that "gorping" meant "ride" than those who heard "The rabbit is gorping the monkey," who were more likely to think it meant "carry." This powerful ability, in which children attend to linguistic context to identify which part of a complex visual scene is labeled by an unfamiliar verb, is known as syntactic

bootstrapping.[1]

However, not all linguistic contexts are equally supportive of verb learning. For example, some contexts that contain useful information may be too difficult to process. Lidz et al. (2009) and He and Lidz (2016) found that 1-year-olds struggled to learn intransitive verbs that were preceded by a lexical content noun (e.g., "The boy is gorping"), but they did better when verbs were preceded by a pronoun (e.g., "He is gorping"). The authors reasoned that the content noun created a processing burden that left children with insufficient resources to learn the novel verb. Extending this finding, He, Kon, and Arunachalam (2020) found that slightly older children, ages 2 and 3 years, overcame this difficulty; they could learn a novel verb when the sentence contained a lexical content noun (e.g., "The boy is gorping"). However, they struggled to learn the verb when the content noun was modified by an adjective (e.g., "The tall boy is gorping"). Taken together, these results suggest that learners struggle with linguistic contexts that impose too great a processing load, and they also suggest that as children's processing capacities develop, they are able to learn from a larger variety of linguistic contexts.

On the other hand, some contexts may be easy to process but too sparse in information for children to benefit from them. Imai et al. (2005; Imai et al., 2008) and Arunachalam and Waxman (2011, 2015) found that English-acquiring preschoolers performed better with more contentful linguistic contexts than less contentful contexts. In Arunachalam and Waxman (2015), for example, 2-year-old children struggled to learn transitive verbs when pronouns flanked the verb (e.g., "Look! He is gorping it") but they did much better when provided with lexical content

[1] Despite the name, this mechanism includes children's use of semantic and referential information as well as syntactic information (e.g., Arunachalam, Syrett, & Chen, 2016; Fisher et al., 1994; Syrett, Arunachalam, & Waxman, 2014).

nouns (e.g., “Look! The boy is gorping the balloon”). Echoing the developmental trajectory in the previous paragraph, Imai et al. (2005; Imai et al., 2008) found that 3-year-olds could succeed with pronouns, but struggled with even less informative contexts in which the subject and object were omitted (e.g., “Look! Gorping!”). These studies suggest that to determine the meaning of a novel verb, children require a certain amount of semantic support (see also Syrett, Arunachalam, & Waxman, 2014), and again, that children become more skilled, requiring less semantic support, with age.

Taken together, this research indicates that the optimal contexts for verb learning are both relatively easy to process and semantically informative. However, these prior studies have only examined children’s learning when they are provided exposure to a novel verb in one linguistic context (e.g., Arunachalam et al., 2016; Fisher et al, 1994; Imai et al., 2005; Imai et al., 2008). Children often hear the same words repeated in several sentences in a short span of time, and those sentences may either be repetitions of the same sentence or variants on the sentence (called variation sets; e.g., Onnis et al., 2008). Certainly in language intervention, speech-language pathologists aim to present a new word multiple times in a single session (e.g., Alt et al., 2020). In considering multiple exposures to a novel verb, we are faced with a new question: Is it better to provide consistent, unvaried input across each exposure, or is it better for there to be some variability across exemplars?

The benefits of variability in word learning

Variability in a word-learning situation may be achieved through visual differences, linguistic differences, or both. In the current study, we are specifically interested in the contrast between “consistent” linguistic input, in which children hear one type of linguistic context repeatedly, with “variable” linguistic input, in which children hear two types of linguistic

contexts—one that is easy to process (but may not contain much information) and another that is more semantically informative (but may be more difficult to process). We hypothesize that providing children with both linguistic contexts will be better than providing them with only one.

Although this particular manipulation of linguistic variability has not yet been studied, there are many reasons to believe that variable presentation of a new vocabulary item in multiple linguistic contexts will be best. Broadly, variability is thought to benefit language learning, including specifically word learning. Visual variability allows children to compare across multiple exemplars and identify commonalities (Childers & Tomasello, 2001; Gentner & Namy, 2006; Perry et al., 2010; Waxman & Klibanoff, 2000; although, see Owen Van Horne & Strother-Garcia, 2020). For example, a child who views multiple kinds of dogs, each paired with the label "dog," should be more likely to correctly generalize the word "dog" to the category of dogs than a child who sees only one dog. For verb learning specifically, the literature contains some conflicting results; some studies find that visual variability in verb learning is helpful, just as it is for noun learning (Gentner & Borodistky, 2001; Twomey et al., 2014; Piccin & Waxman, 2007). Other studies find better learning when children view only one agent executing the referent action as compared to multiple agents (Childers et al., 2017; Maguire et al., 2008). Snape and Krott (2018) found that for learning transitive verbs, children performed best when provided scenes with variable patient/theme objects but a consistent agent. Relatedly, in a preposition learning task, Nicholas, Alt, and Hauwiller (2019) find complex patterns wherein the ability to benefit from variability in the objects with which prepositions are presented depends on children's receptive language abilities; preposition learning is similar to verb learning in that it requires noting relations among multiple entities. Thus, the role of variability in visual exemplars in verb learning tasks, and in word learning generally, is nuanced (for a review, see Horvath &

Arunachalam, 2019).

When it comes to the linguistic context in which the verb appears, there is evidence that variability is useful for learners. This work has focused on sentences that differ in their argument structure. For example, verbs that denote a change of state (e.g., *melt*) or caused motion (e.g., *bounce*) typically can appear in both transitive and intransitive structures in the causative alternation (*The sun melted the ice / The ice melted; The basketball player bounced the ball / The ball bounced*), while verbs that denote an action upon an object without encoding a necessary change in that object (e.g., *kick, push*) do not alternate (*The girl kicked the wall / #The wall kicked; The cat pushed the table / #The table pushed*). By two years of age, children can take advantage of these patterns to learn new verbs; children hearing a verb in both structures of the causative alternation are more likely to assign it a change of state meaning than those who hear just one of the structures (Bunger & Lidz, 2004; Naigles, 1996; Scott & Fisher, 2009).

The linguistic variability we examine in the current study, however, does not involve syntactic structure. Instead, we compare two linguistic contexts that differ in how the subject and object are realized, specifically in terms of how much semantic content they provide—as content nouns or pronouns. We chose this particular manipulation for several reasons. First, alternation between content nouns and pronouns is particularly frequent in speech; the first mention of a referent is likely to involve a content noun, but speakers subsequently shift to a pronoun for that same referent. Consider, for example, this excerpt of caregiver speech from Suppes (1974): “The dolly will ride on the bicycle…. She’s riding on the bicycle” (filename “nina05.cha” in the CHILDES database, MacWhinney, 2000).

Second, pronouns and content nouns may each support different aspects of the verb-learning process. As we have noted above, in verb-learning tasks, pronouns appear to be helpful

in that they may impose less of a processing load on young learners than lexical content nouns. Pronouns can also be helpful for the initial tasks of parsing a word form from the ongoing speech stream and determining its grammatical category (i.e., whether it is a noun or a verb) (e.g., Mintz, 2003). Because pronouns are so frequent, if a novel word is flanked by two pronouns (e.g., "He's gorping it"), children can easily identify the phonological boundaries of the novel word form. Similarly, this "frequent frame" can help children determine that the word is a verb because they have had substantial experience hearing verbs flanked by pronouns (e.g., Cauvet et al., 2014; Frost et al., 2019; Mintz, 2003, 2006; Shi & Melançon, 2010).

However, as we have also seen, for the more complex task of mapping verbs to meaning, the semantic content of content nouns can be more helpful than the sparse semantic content of pronouns. Therefore, given prior findings that linguistic variability supports verb learning (e.g., Naigles, 1996), and the fact that content nouns and pronouns each support different aspects of the verb learning task, we hypothesize that children will perform better provided some sentences with content nouns and some with pronouns (hereafter the "Varied condition") as compared to only hearing sentences with repeated content nouns ("Consistent condition").

In support of this hypothesis, Childers and Tomasello (2001) found that variability with content nouns and pronouns supported children's abilities to extend a novel verb from the syntactic frame in which they had heard it to a new frame. Children were more successful at producing a novel verb in a transitive frame, and at following an instruction given in a transitive frame, if they had heard that verb in an intransitive frame with both content nouns and pronouns than if they had heard it with only content nouns. However, the training that Childers and Tomasello provided involved multiple different kinds of agents and patients, resulting in linguistic variability in both the content nouns and the pronouns used in both conditions. Thus,

even the condition that only provided content nouns offered variability, both visually and in semantic content.

Following closely on the designs used by Imai et al. (2005; Imai et al., 2008) and Arunachalam and Waxman (2011, 2015), in the current study we first presented children with a novel verb (e.g., "ziff") as they viewed a corresponding event (e.g., a girl lifting and lowering a truck). In the Varied condition, the novel verb occurred both in transitive sentences with content nouns (e.g., "The girl is ziffing the truck") and in transitive sentences with pronouns (e.g., "She is ziffing it"). In the Consistent condition, the novel verb only occurred in sentences with content nouns. We did not include a condition of consistent pronoun use because previous studies have shown that in this design children of this age do not learn novel transitive verbs from pronominal contexts alone (e.g., Arunachalam & Waxman, 2011, 2015). Next, we tested whether children learned the verb's meaning by presenting two new scenes side-by-side, one depicting the agent performing a different action on the object (e.g., the girl dumping the truck bed), and the other depicting the same action but on a different object (e.g., the girl lifting and lowering a teddy bear).

Typically developing versus late-talking toddlers

Most word learning research has looked at typically developing children as a homogeneous group, despite large differences in individual language ability. Learning verb meanings by attending to linguistic context relies on a host of other abilities, including having some baseline level of grammatical and vocabulary knowledge, and practiced parsing skills that allow children to rapidly comprehend the linguistic context. It is therefore likely that children differ in how and how well they use linguistic context in verb learning. For this reason, in the current study we asked whether performance in the Varied and Consistent conditions varies by

language ability—specifically, contrasting typically developing and late-talking toddlers. Late talkers are children with atypically small expressive vocabularies with no known cause—they have age-appropriate motor skills, social skills, and play skills, and no diagnoses expected to affect cognition. Some may also have receptive language delays, but others appear to have typical receptive language skills. Further, late talkers appear to differ not only in overall vocabulary size but also in vocabulary composition (Horvath, Rescorla, & Arunachalam, 2019; MacRoy-Higgins et al., 2016; Rescorla et al., 2001) and structure (Beckage et al., 2011; c.f., Jimenez & Hills, 2017). Approximately 15% of 2-year-olds are late talkers, but the majority of late talkers develop language skills within the typical range within a few years (Desmarais et al., 2008; Rescorla, 1989).

Few studies have considered how late talkers acquire vocabulary, and, to date, none of these has examined verb learning. However, studies on noun learning suggest that late talkers are not as adept at using the same word-learning strategies as their typically developing peers. For example, late talkers are not as successful in fast-mapping novel nouns as typically developing children (Ellis Weismer & Evans, 2002; Ellis Weismer et al., 2013). Research by Colunga and others has also found that late talkers may not reliably use "shape bias" as a cue for noun learning (e.g., Beckage & Colunga, 2019; Colunga & Sims, 2017; Jones, 2003; Sims et al., 2016). It is likely that these differences extend to verb learning; late talkers differ in the types of verbs they acquire as compared to typically developing peers (Horvath et al., in revision; Horvath et al., 2019; Jimenez et al., 2020). This may be in part because of processing demands; we have seen that children must be able to process the linguistic context in order to benefit from it in verb learning (e.g., He et al., 2020), but late talkers are slower to process language than typically developing children (Fernald & Marchman, 2012).

Therefore, we hypothesize that late talkers might not be able to benefit from variability and might perform better in the Consistent condition than the Varied condition. Repetition can support word learning in children with language delays and disorders (e.g., Rice et al., 1994; Riches et al., 2005), and children who are slower language processors benefit from quantity of exposures more than diversity of exposures as compared to children who are faster processors (Jones & Rowland, 2017).

To summarize, two research questions were addressed in the current study: First, do children show better acquisition of novel verbs when presented in two different linguistic contexts (Varied Condition), or in the same linguistic context repeatedly (Consistent Condition)? We hypothesize that children will be better able to learn novel verbs provided variable linguistic contexts. Second, does the answer to the first question depend on the child's expressive vocabulary (late talkers as compared to typically developing children)? We hypothesize that late talkers will struggle with the processing demands of variability and instead perform better when provided consistent input.

Methods

Participants

The final sample included 47 English-acquiring 2-year-olds (18 females, 29 males). The gender distribution was skewed toward males because we oversampled late talkers, who are more likely to be male (Scheffner Hammer et al., 2017; Zubrick et al., 2007). Children ranged in age from 24.5 to 35.8 months (M = 28.5 months). Recruitment and testing procedures were approved by Boston University's Institutional Review Board, and parents provided written consent on behalf of their children.

Participants were reported by their parents to be exposed to English at least 80% of the time and to have no developmental disorders other than suspected language delay; four children were reported to have "mild language delay," "language delay," or "expressive language delay." Because delayed language is also common in children with autism spectrum disorder, participants were screened using the Modified Checklist for Autism in Toddlers, Revised (M-CHAT-R; Robins et al., 2009), a parent report screener intended to identify autism risk. All included children received a score indicating "low risk" for autism spectrum disorder.

Parents provided demographic information. The sample was 86% White, 4% Asian and 4% from one or more races; 2 families declined to provide racial information. One child was reported to be Hispanic/Latinx. The majority of participants in the sample (77%) had at least one parent with a Master's degree or higher. One child did not have a parent who had completed a Bachelor's degree; four families (9%) did not provide education information.

Nine additional children participated in the study but were excluded from the final sample. Four were excluded because of developmental concerns that might affect language development beyond being late talkers: One received a diagnosis of autism spectrum disorder just after participation, one had a history of tongue-tie (ankyloglossia), and two had a history of ear tubes. Five additional children were excluded because they contributed insufficient eye-tracking data (see below).

Parents provided information about children's vocabulary size using the MacArthur-Bates Communicative Development Inventories Level 2 Short Form A (MCDI: Fenson et al., 2000). Scores ranged from 1 to 100 (M = 68, SD = 28). We classified children as late talkers in two ways. First, children whose MCDI score was at or below the 15th percentile criterion for their age and gender were categorized as late talkers, as in many prior studies (e.g., Dale et al.,

2003; Ellis et al., 2015; MacRoy-Higgins et al., 2013).[2] The MCDI is only normed for children up to age 30 months; we used the 30-month percentile data for children older than 30 months; two children were identified as late talkers through this method. Second, we classed two additional participants (aged 32 and 35 months) as late talkers because their parent reported that they had received speech-language therapy for language delay. With these criteria applied, a total of 14 children in the sample were classed as late talkers, and the remaining 33 as typically developing. Late talkers averaged significantly lower scores on the MCDI ($t = 7.83$, $p < 0.001$). In raw numbers, there was a clear split in MCDI words produced; of the late talkers, only the two who were identified for their history of speech-language therapy produced more than 50 of the MBCI words. All but two typically developing toddlers (both male, ages 25 and 26 months) produced at least 50 of the MCDI words.

[2] Because the term "late talker" is not a clinical diagnosis, prior research has varied in the criteria used to identify this subgroup. Late talkers are classified based on the size of their expressive vocabulary, with individual studies using a cutoff point somewhere between the 10th percentile (e.g., Bishop et al., 2012) and 25th percentile (e.g., Colunga & Sims, 2017), with many studies using the 15th percentile (Dale et al., 2003; Ellis et al., 2015; MacRoy-Higgins et al., 2013). We use the 15th percentile as a compromise and because it is the closest cutoff to estimated population rates, which typically range from 13 to 15% of toddlers (Desmarais et al., 2008; Zubrick et al., 2007). We note here that although our a priori cutoff was the 15th percentile, all late talkers in our sample meet the more conservative threshold of the 10th percentile, except for the two who were classified based on their history of speech-language therapy. We also note that three typically developing toddlers (9% of the sample) had MCDI scores between the 20th and 30th percentiles.

The first author, a licensed speech-language pathologist, also administered the Preschool Language Scales, 5th edition (Zimmerman et al., 2011) for a fuller picture of children's language comprehension and production. There were significant group differences on both the Expressive Communication subscale ($t = 5.15$, $p < .001$) and Auditory Comprehension subscale ($t = 6.15$, $p < 0.001$). See Table 1. There were no differences between the two groups with respect to average age ($t = 1.60$, $p = .11$, *n.s.*) or proportion male ($z = 0.89$, $p = .37$, *n.s.*). There was no significant difference between groups in the proportion of participants whose parents had a postgraduate degree ($z = 1.75$, $p = 0.08$, *n.s.*).

--- TABLE 1 ABOUT HERE---

Apparatus

The experimental task was presented on a 24-inch Tobii T60 XL corneal reflection eye-tracker monitor, which samples at 60 Hz, running Tobii Studio software. Children sat in front of the monitor either in a car seat or on their parent's lap. If the latter, the parent wore a blindfold.

---TABLE 2 ABOUT HERE---

Materials

For the visual stimuli, we recorded video clips of actors acting on objects. See Table 2 for a description of visual stimuli on each trial. Each trial included three scenes, one used to familiarize children to the novel verb and two used in the test phase. The actor was consistent among all three scenes in each trial. There was no repetition of actors or objects across trials.

For the auditory stimuli, a female speaker of American English recorded sentences using a child-directed speech register in a sound-attenuated booth. We then edited the visual and auditory stimuli in Final Cut Pro software to create the trial structure shown in Figure 1. Stimuli were very similar to those used in prior verb learning studies by Arunachalam and Waxman

(2010, 2015) except that the auditory stimuli included multiple presentations of the novel verb. In the Consistent Condition, all sentences contained content nouns with definite objects (e.g., "the girl," "the truck"). In the Varied Condition, half of the sentences contained content nouns with definite objects, and half of the sentences contained only pronouns (e.g., "she," "it"). Both conditions included variations of tense/aspect marking on the verb (e.g., "The girl is ziffing the truck! Look! The girl ziffed the truck."), which was done for pragmatic felicity, as the visual scenes were described before, during, and after the events occurred. However, tense/aspect marking was identical across conditions and trials. See Figure 1. Children additionally heard directives to find the target in the gerund form (e.g., "Let's find ziffing!", "Where is ziffing?"). The novel verbs are listed in Table 2.[3]

Design

Children participated in a verb learning paradigm consisting of eight experimental trials (though only seven were analyzed, see below) in a within-subject design. See Figure 1. Each trial consisted of four phases: Familiarization, Preview, Prompt, and Test. During the Familiarization phase, which lasted approximately 30 seconds, children first viewed a still frame of an actor holding an object (e.g., a woman holding a toy dump truck), then viewed a dynamic scene (e.g., the woman lifting and lowering the truck). This repeated twice. The accompanying auditory

[3] Each verb form occurred in both the Consistent and Varied conditions, counterbalanced across participants, so phoneme probability of the verb stems should not affect interpretation of our results. However, we note that phoneme probabilities were similar across the novel words; the sum of phoneme probabilities for the verb stems, calculated using the Phonotactic Probability Calculator (Vitevitch & Luce, 2004), ranged from 0.07 to 0.29 (nork = 0.16, pell = 0.23, pilk = 0.29, sem = 0.22, tope = 0.13, wug = 0.07, zif = 0.12).

stimuli introduced a novel verb (e.g., "ziff"), which children heard eight times. In the Consistent Condition, the novel verb was always flanked by the same two content nouns which labeled the event participants (e.g., "The girl is gonna ziff the truck"). In the Varied Condition, auditory stimuli began with a sentence introducing the actor and object (e.g., "Let's see a girl and a truck") consistent with Arunachalam and Waxman (2015). The novel verb was flanked by content nouns in half of the sentences, but the other half contained pronouns (e.g., "She is gonna ziff it"). Visual scenes were always paired with the same verb, but were counterbalanced to appear in both the Consistent and Varied conditions.

The three phases following the Familiarization phase (Preview, Prompt, and Test) were identical across conditions. During the Preview phase (6 seconds), two new dynamic scenes played side by side. In one, the same actor acted on the same object, but with a different action; in the other, the actor performed the familiar action, but on a different object. During this phase, the two scenes played with attention-getting audio (e.g., "Look!") but no novel words; this phase was designed simply to allow participants to observe the two scenes.

Next, during the Prompt phase (2 seconds), the scenes disappeared, replaced by a centrally positioned yellow star to direct visual attention to the center of the screen. The audio prompted a search for the novel verb's referent (e.g., "Let's find ziffing!").

Immediately afterward was the Test phase (6 seconds). In this phase, the two test scenes reappeared in their original locations with another audio prompt (e.g., "Where is ziffing?").

---FIGURE 1 ABOUT HERE---

On all trials, the scene depicting the familiarized action was the target scene. To succeed, children had to attend to the audio presented during Familiarization, assign the novel verb an action meaning, and extend the verb to apply to a new scene in the Test phase. Prior work shows

that 2-year-olds perform better in this task when the novel verb is presented with content nouns than with pronouns (Arunachalam & Waxman, 2011, 2015). Manipulating the audio presented during Familiarization allowed us to determine whether the Consistent condition (with only content noun contexts) or Varied condition (with both pronominal and content noun contexts) better supports learning.

Trials were presented in blocks, with the order of the Consistent and Varied conditions counterbalanced across participants, and the condition to which each trial was assigned also counterbalanced across participants. The blocks were separated by an 11-second break in which animated shapes moved across the screen accompanied by a clip of instrumental music.

Procedure

Children participated in the experimental task as part of a 2-visit protocol. At the first visit, children first played with an experimenter while parents provided informed consent and completed a demographic questionnaire, the MCDI, and the M-CHAT-R. Children then participated in an unrelated experimental task and the Preschool Language Scales. At the second visit, approximately two weeks later, children again began the visit by playing with an experimenter. They then entered the testing room where they were seated in front of the eye-tracker monitor and viewed the experimental task.

Data Processing and Analysis Plan

To evaluate children's learning in each condition, we evaluated their gaze behavior during the Test phase of each trial. Following prior work, our first planned analysis focused on the time window from 1 to 2.5 seconds of the Test phase. Recall that, in the Prompt phase immediately prior, children heard a directive to find the target while their visual attention was directed to the center of the screen with a central fixation star. Prior novel verb-learning studies

with this design have found that 2-year-olds require approximately 1 second to disengage attention from the centrally-placed star and to program and launch an eye movement in response to the auditory prompt they have just heard (e.g., Arunachalam, 2013; Arunachalam & Dennis, 2018). The goal of this analysis was to determine whether there were differences in preference for the target scene across the Consistent and Varied conditions, and whether these differences were related to children's expressive vocabulary size. Because we did not have an equal number of typically developing and late talking children, our primary analysis treated expressive vocabulary as a continuous measure rather than grouping children. In the second analysis described below, we then used a binary categorization of typically developing vs. late talking.

We used mixed-effects linear modeling (with the lme4 package version 1.1-12 (Bates et al., 2015) and lmerTest package version 2.0-36 (Kuznetsova et al., 2017) in R version 3.3.0 (R Core Team, 2016). The dependent variable was the proportion of looks to the target scene (i.e., the preserved action scene) versus elsewhere in the 1- to 2-.5-second window of the Test phase; note that the denominator for the analysis includes looks to neither scene and track loss, which is a more conservative approach than excluding these data points, and it has been used in prior eye-tracking work with similar paradigms (e.g., Borovsky & Peters, 2019; Horvath et al., 2018). Following Barr (2008), the gaze data was collapsed in 50-ms bins to reduce the effects of eye movement-based dependencies (that is, the direction of gaze at one time point is not independent of direction of gaze at the next), and the proportions were empirical logit transformed. The regression models included the random effects structure of a by-participant intercept with a slope for time, and a by-trial intercept with a slope for time; these random effects are important because different children will shift gaze at different speeds, and the different video scenes used in different trials may also yield differences in how quickly children visually scan them. In the

first model, meant to examine the relationship between vocabulary and performance, we included fixed effects of time (which we include because children's gaze will vary across the test phase), expressive vocabulary score on the MCDI (from 1-100) and condition (contrast coded as 0.5 for the Consistent condition and -0.5 for the Varied condition), along with their interactions. We also included age in months (centered around its mean) as a fixed effect. Models were fit using weighted restricted effects maximum likelihood. The second model directly tested the hypothesis that late talkers might be different from their typically developing peers. This model included the same outcome variable and random effects structure as the first. The fixed effects for the model were time, condition, and group (late talker and typically developing). A third, post-hoc analysis asked how many children showed a pattern of better learning in one condition as compared to the other.

We first processed the data to evaluate tracking quality. In doing so, we realized that one of the eight items depicted a motion event in which the actor was not in the visual scene at the very beginning of the scene (an actor pulling a toy drum across the floor); we excluded this item from analysis because, unsurprisingly, children's gaze was not directed to this scene early in each phase. Because we counterbalanced the assignment of items to condition, this did not affect the amount of data in each condition.

We also conducted a preliminary analysis to determine if there was an effect of the order of the condition blocks (Consistent trials first vs. Varied trials first). There was none, so we collapsed across the two orders for the remainder of the analyses.

Exclusionary criteria

Children who had greater track loss for greater than 50% of the test phase on more than half of their trials and children who did not contribute at least one trial in each condition were

417 excluded (n = 5 participants from initial 56 recruited). We chose this threshold on the basis of
418 prior studies using a very similar paradigm (e.g., Arunachalam & Dennis, 2018; Horvath et al.,
419 2018). For the remaining children in the sample, all trials were included. There was a small non-
420 significant negative correlation between vocabulary size on the MCDI and percentage of track
421 loss for the included participants (R = -0.15, p = .40) such that lower vocabulary was associated
422 with greater track loss. Being a late talker may be associated with behavioral and attentional
423 differences (e.g., Carson et al., 1998; Horwitz et al., 2003; Irwin et al., 2002), so this is not
424 surprising.
425
426 **Results**
427 As shown in Figure 2, participants looked more at the target scene in the Consistent
428 condition than the Varied condition. Recall that the denominator of the proportion of target
429 looking includes looks to neither scene and track loss, so 0.50 should not be interpreted as
430 chance looking. The primary analysis of looking to the target scene as a function of time,
431 condition, and expressive vocabulary size, as well as their interactions, revealed a significant
432 main effect of condition, and significant interactions between time and condition, condition and
433 vocabulary size, and a significant three-way interaction of time by condition by vocabulary size.
434 See Table 3. The main effect of condition reveals that participants looked to the target scene
435 significantly more in the Consistent condition (M = 0.42, SD = 0.20) than the Varied condition
436 (M = 0.37, SD = 0.19). The interaction between condition and time indicates that children's gaze
437 patterns toward the target scene also vary across the analysis window as a function of condition.
438 The interaction between condition and vocabulary reveals that the effect of condition is more
439 pronounced for children with lower vocabularies than it is for those with higher vocabularies.

Finally, there is a significant three-way interaction of time by condition by vocabulary: The impact of condition on gaze behavior across time during the test window is itself impacted by children's overall vocabulary size, indicating complex relations among these factors. The model also revealed a significant fixed effect of age, wherein younger children looked more to the target scene overall than did older children.

---FIGURE 2 ABOUT HERE---

---TABLE 3 ABOUT HERE---

The second analysis, in which we grouped children as either late talkers or typically developing, yielded the same pattern of effects. See Figures 3A and 3B. The regression indicated no main effect of Group; however, there was a significant interaction between Condition and Group. See Table 4. To understand this interaction, we conducted a post-hoc analysis of estimated marginal means using the emmeans package, version 1.3.3 (Lenth, 2019); *p*-values were adjusted using the Tukey post-hoc method. The results revealed that both late talkers and typically developing children performed significantly better in the Consistent condition than in the Varied condition, but that this difference was more pronounced for late talkers (Consistent: $M = 0.40$, $SD = 0.20$; Varied: $M = 0.36$, $SD = 0.19$; estimate = 0.26; $SE = -.06$, $p < .001$) than for typically developing children (Consistent: $M = 0.43$, $SD = 0.20$; Varied: $M = 0.38$, $SD = 0.20$; estimate = 0.20, $SE = 0.04$, $p < .001$). Because late talkers differ in whether their receptive language is also delayed, and this factor may critically distinguish those who go on to have language problems and those who do not, we also added the receptive language measure of standard score on the Auditory Comprehension subscale of the Preschool Language Scales, but here too, the same pattern held and there was no significant effect of Auditory Comprehension

score (p = .14).[4]

---FIGURE 3A ABOUT HERE---

---FIGURE 3B ABOUT HERE---

Finally, we conducted a post-hoc assessment of individual patterns to see how widespread the pattern was of learning better in the Consistent condition as compared to the Varied condition. We compared each child's proportion of looking to the target scene during the target window across the two conditions. Unsurprisingly, the majority of children looked to the target more in the Consistent condition (n = 29) than the Varied condition (n = 18). However, there were no differences in the proportion of time children spent looking to the target scene in their better learning condition (the Consistent condition: M = 0.51, SD = 0.17; the Varied condition: M = 0.49, SD = 0.17; t = 0.39, p = 0.70, *n.s.*).

---TABLE 4 ABOUT HERE---

[4] We ran an additional post-hoc analysis to determine whether receptive language skills alone predicted performance. The model included the dependent variable of proportion of looks to the target scene, the random effects of participant (with a slope for time) and trial (with a slope for time), and the fixed effects of time, condition, and Preschool Language Scales Auditory Comprehension Standard Score, along with their interactions. The regression revealed no main effect of receptive language skills (b = 0.019, SE = 0.013, t = 1.40, p = .16). This is not altogether surprising: the Auditory Comprehension subscale captures broad receptive language abilities, including following directions, understanding questions, and receptive vocabulary. Our task, though receptive, focuses specifically on vocabulary acquisition, and as such is more likely to be predicted by concurrent vocabulary measures (both expressive and receptive).

Discussion

Although verb learning is a notoriously difficult task, it has been robustly demonstrated that children are sensitive to the linguistic context in which an unfamiliar verb appears and can use it to acquire verb meanings (e.g., Naigles, 1990; Fisher, 2002). However, we still know little about what kinds of linguistic contexts are best. Given that learners often hear a single word multiple times within a conversation, and that speech-language pathologists carefully plan how and how many times they will introduce a new word in a session, we sought to learn more about how multiple kinds of verb learning contexts would compare to repetition of a single type of context.

We presented children with novel verbs either in linguistic contexts containing only content nouns (Consistent condition) or in a combination of contexts containing content nouns and contexts containing pronouns (Varied condition). Although variability has been shown to support many aspects of language learning, our findings do not support the hypothesis that varied presentation in both types of linguistic contexts is better than consistent presentation in one type of context for acquiring verb meaning.

Why wasn't variability more helpful for learners than repetition, given prior work illustrating benefits of variability in other word-learning tasks? One likely possibility is that the content noun context better supported children's abilities to establish a representation of the novel verb's meaning, and that repetition of the verb in the same context solidified that representation for the verb, making it easier to access at test. Consider Arunachalam and Waxman (2015)'s finding that typically developing children are unable to resolve pronouns with their referents in order to acquire a novel verb's meaning (e.g., "He is going to blick it"), even when there were only two referents in the visual scene. This was true even when children were

provided the content nouns in a previous sentence (e.g., “Let’s see a boy and a balloon. He is going to blick it”). This added step of resolving the pronouns with their referents may be too cognitively demanding, leaving children insufficient resources to acquire the verb. This interpretation is consistent with work by Hadley et al. (Hadley & Walsh, 2014; Hadley et al., 2017), who have demonstrated the benefits of using content nouns in subject position. Content nouns, Hadley et al. argue, help support verb acquisition in particular because they better highlight the relationship among event participants than do pronouns. (Hadley et al.’s studies do not contrast consistency and variability, but rather content nouns versus pronouns. In both instances, they included variability in the subject position.) Under this interpretation, children in our study were able to solidify their representations of the novel verb across all of the eight mentions of the verb during the familiarization phase in the Consistent condition, but were only able to access four of the eight mentions of the verb in the Varied condition. If instead the task of parsing out the phonological form of the novel verb, or determining that it was a verb rather than a noun, had been more challenging for children at this age, the pronominal context may have provided much needed support for getting the learning task off the ground. (Younger children are indeed likely to require such support; He & Lidz, 2016).

Another possibility is that variability in linguistic contexts is simply not useful for verb learning unless the syntactic frame is manipulated in such a way that it reveals something about the verb’s semantic class. For example, hearing a verb in the causative alternation (e.g., “The ball is blicking” / “The baby is blicking the ball”) offers different information from hearing a verb in the implicit object alternation (e.g., “The baby is blicking” / “The baby is blicking a cookie”). Naigles (1996) and Scott and Fisher (2009) demonstrated that, when each context provides different information, the learner derives benefit from this type of variability (see also Bunger &

Lidz, 2004). By contrast, the two frames presented in the current study do not provide different kinds of information about the verb's meaning.

Although our data is compatible with this second possibility, we think the first possibility is more likely given that pronominal contexts have been shown to play an important role in some aspects of language learning (e.g., Childers & Tomasello, 2001) and given that a large body of work documents a trade-off between how informative the linguistic context is and how easy it is to process for acquiring verb meaning (see, e.g., He & Arunachalam, 2017 for review).

Our findings are particularly interesting given that the linguistic pattern of the Varied condition, in which content nouns were replaced with pronouns after an initial use, is thought to be more felicitous in discourse (e.g., Gordon et al., 1993) and adults, at least, incur a processing penalty when names are repeated (e.g., Gordon et al., 1993; Gordon & Scearce, 1995). However, repeated names (as opposed to pronouns) may help to enhance memory representations (e.g., Gernsbacher, 1989) and therefore, for 2-year-olds solving the task of acquiring a novel verb, repetition of content nouns may be a boon rather than a hindrance.

In the long term, however, variability in how verbs are presented, as we would expect in a natural learning environment, very likely supports learning and generalization (e.g., Hoff-Ginsberg, 1985; Naigles & Hoff-Ginsberg, 1995, 1998; but see also Hsu, Hadley, & Rispoli, 2017). The current results are relevant for our understanding of fast mapping and of children's initial efforts to acquire new verbs. Subsequent extension of that verb into new contexts, once a sufficiently robust representation is established, may be dependent on variability in both the linguistic and visual contexts in which it occurs.

Another important finding from this work is that late talkers and typically developing toddlers differed in their abilities to learn the novel verb. In the Consistent condition, there was

no difference between groups in performance; however, in the Varied condition, late talkers performed substantially worse than typically developing children. We offer two interpretations, not incompatible, for this difference. The first interpretation is that this difference is due to the poorer language processing abilities of late talkers. Recall that in the Varied condition, children heard the novel verb four times, each time flanked by pronouns. Resolving the referents of pronouns requires language processing skill and skill at integrating information from different sources (e.g., the visual scene and the linguistic context), and there is some evidence from older children that those with poorer language comprehension skill may have difficulty resolving pronoun referents (e.g., Oakhill & Yuill, 1986).

Additionally or alternatively, late talkers may require more exposure to a novel verb in informative linguistic contexts than their typically developing peers. If late talkers learn little or nothing from pronoun-only contexts, then in the Varied condition, they had access to only four useful exposures, which may be too few for this group. This is consistent with prior work which indicates that children with processing deficits typically require a greater number of exposures in order to learn word meanings (Jones & Rowland, 2017).

Limitations

One limitation of the current study is that because the task was nearly identical to prior studies, we did not include a baseline, no-word condition as a comparison point for measuring learning. Although children preview the test scenes before they are asked to find the target word, this preview comes after children have been exposed to the novel verb, and thus children's looking behavior during this preview is likely influenced by the linguistic input of the familiarization phase. We are therefore unable to make direct comparisons between preview and test or to measure learning by comparing the two. However, our goal was not to demonstrate that

children can learn in one condition and cannot learn in another—we took as precedent the robust evidence from prior work that 2-year-olds can learn a novel verb flanked by content nouns in this task (e.g., Arunachalam & Waxman, 2011, 2015; Waxman et al., 2009). Instead, we wanted to ascertain whether one condition showed better performance than another, which we did find.

A second limitation of this study is our sample, which is not representative. More than three-quarters of the children in our sample had a parent with a graduate degree. Likely as a consequence of this, the typically developing children averaged standard scores on the Preschool Language Scales nearly one full standard deviation above the norm. Even the late talkers, classified based on vocabulary size, performed better than would be expected on this measure. Additionally, more than 80% of our sample was white, which is not reflective of current US demographics.

Finally, we were also limited in our classification of late talkers by our instrument of choice: the MCDI. This is only normed for children up to 30 months, but we recruited children between 24 and 36 months, and we used the 30-month benchmarks for the older half of the sample. However, a 35-month-old who classifies as a "late talker" based on 30-month-old percentiles is more delayed than a 30-month-old who does. This discrepancy may partially explain why we observed an unusual age effect in the first regression model, wherein younger children looked more to the target than older children: Older late talkers, at least, were more severely delayed. Although we added an additional criterion (history of speech therapy) to partially alleviate this problem, we note that it likely still has consequences for the interpretation of our findings.

Conclusions

We observed that in a novel verb learning task, providing children with consistent content

noun use better supports verb learning than does variable content noun and pronoun use. The finding that content nouns strongly support verb learning is consistent with other work highlighting the importance of content nouns in verb learning (e.g., Arunachalam & Waxman, 2015; Hadley & Walsh, 2014).

However, we do not suggest that one verb-learning context is uniformly optimal for all learning tasks or for all children; instead, the usefulness of any given linguistic context in any given situation will vary according to many factors, including age and language ability (Horvath & Arunachalam, 2019). For example, younger children or children with poorer language processing abilities may struggle with the processing demands of a content noun context (He & Lidz, 2016), while older children may succeed (He et al., 2020). The current results suggest that variability, too, may have greater or lesser benefits across age because of the tradeoff of processing and informativity (see also Nicholas et al., 2019 for related results for preposition learning). Different kinds of variability other than the one we manipulated here are also likely to yield different patterns.

Our finding supports the hypothesis that optimal contexts for word learning (and in particular verb learning) will vary as a factor of language ability. This finding is notable given the relatively small literature on word learning in late talkers. Although late talkers are defined by their vocabulary size, relatively few studies have explored how they learn new word meanings, and this research has been limited to learning nouns (e.g., Alt et al., 2020; Capone et al., 2020; Ellis et al., 2015, Ellis Weismer et al., 2013; Jones, 2003; Jones & Smith, 2005; MacRoy-Higgins & Montemarano, 2016). However, late talkers show subtle differences in their verb vocabulary composition (Horvath et al., in revision; Horvath et al., 2019; Jimenez et al., 2020), which may indicate differences in verb-learning processes. The findings from this study

suggest that children with smaller vocabularies—including late talkers—are not as successful in verb learning provided non-optimal linguistic contexts.

This finding may have implications for the development of clinical interventions seeking to teach children new vocabulary. While we are hesitant to make direct clinical recommendations from one experimental task, we advocate for two lines of future research: One focused on better defining optimal verb-learning contexts for learners with diverse abilities, including typically developing children as well as children with or at risk for language disorder, and another translating these basic science findings into clinical intervention. We also advocate for interventions specifically focusing on verb vocabulary because of the critical role that verbs play in the development of grammar (e.g., Ebbels, van der Lely, & Dockrell, 2007; Hadley et al., 2016; Oetting, Rice, & Swank, 1995).

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

Arunachalam, S. (2013). Two-year-olds can begin to acquire verb meanings in socially impoverished contexts. *Cognition, 129*, 569–573. https://doi.org/10.1016/j.cognition.2013.08.021

Arunachalam, S. & Dennis, S. (2018). Semantic detail in the developing verb lexicon: An extension of Naigles and Kako (1993). *Developmental Science, 22*(1), e12697. https://doi.org/10.1111/desc.12697

Arunachalam, S., & He, A. X. (2018). Children's acquisition of nouns that denote events. In *Proceedings of the 42nd Boston University Conference on Language Development (BUCLD)* (pp. 29-44). Somervillle, MA: Cascadilla Press.

Arunachalam, S., Syrett, K., & Chen, Y. (2016). Lexical disambiguation in verb learning: Evidence from the conjoined-subject intransitive frame in English and Mandarin Chinese. *Frontiers in psychology*, *7*, 138. https://doi.org/10.3389/fpsyg.2016.00138

Arunachalam, S., & Waxman, S. R. (2010). Specifying the role of linguistic information in verb learning. In *Proceedings of the 34th annual Boston University conference on language development* (pp. 11-21). Somerville, MA: Cascadilla Press.

Arunachalam, S., & Waxman, S. R. (2011). Grammatical form and semantic context in verb learning. *Language Learning and Development*, *7*(3), 169–184. https://doi.org/10.1080/15475441.2011.573760

Arunachalam, S., & Waxman, S. R. (2015). Let's see a boy and a balloon: Argument labels and syntactic frame in verb learning. *Language Acquisition*, *22*(2), 117– 131. https://doi.org/10.1080/10489223.2014.928300

Barr, D. J. (2008). Analyzing "visual world" eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*(4), 457–474. https://doi.org/10.1016/j.jml.2007.09.002

Bates, D., Maechler, 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

Beckage, N. M., & Colunga, E. (2019). Network growth modeling to capture individual lexical learning. *Complexity*, *2019*. https://doi.org/10.1155/2019/7690869

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

Borovsky, A., & Peters, R. E. (2019). Vocabulary size and structure affects real-time lexical recognition in 18-month-olds. *PloS one*, *14*(7). https://doi.org/10.1371/journal.pone.0219290

Bunger, A., & Lidz, J. (2004). Syntactic bootstrapping and the internal structure of causative events. In *Proceedings of the 28th Annual Boston University Conference on Language Development* (Vol. 28, pp. 74-85). Somerville, MA: Cascadilla Press.

Capone Singleton, N., & Anderson, L. (2020). Making Object Shape Explicit for Toddlers With Late Language Emergence. *Journal of Speech, Language, and Hearing Research*, *63*(3), 749-763. https://doi.org/10.1044/2019_JSLHR-19-00235

Carson, D. K., Klee, T., Perry, C. K., Muskina, G., & Donaghy, T. (1998). Comparisons of children with delayed and normal language at 24 months of age on measures of behavioral difficulties, social and cognitive development. *Infant Mental Health Journal: Official Publication of The World Association for Infant Mental Health*, *19*(1), 59-75.

Cauvet, E., Limissuri, R., Millotte, S., Skoruppa, K., Cabrol, D., & Christophe, A. (2014). Function words constrain on-line recognition of verbs and nouns in French 18-month-olds. *Language Learning and Development*, *10*(1), 1-18.

Childers, J. B., Paik, J. H., Flores, M., Lai, G., & Dolan, M. (2017). Does variability across events affect verb learning in English, Mandarin, and Korean? *Cognitive Science, 41*, 808–830. https://doi.org/10.1111/cogs.12398

Childers, J. B., & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology, 37,* 739-748. https://doi.org/1037//0012-1649.37.6.73

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

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*,544–560. https://doi.org/ 10.1044/1092-4388(2003/044)

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

Ebbels, S. H., Van Der Lely, H. K., & Dockrell, J. E. (2007). Intervention for verb argument structure in children with persistent SLI: A randomized control trial. *Journal of Speech, Language, and Hearing Research, 50*(5), 1330-1349. https://doi.org/10.1044/1092-4388(2007/093

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

Ellis Weismer, S., & Evans, J. L. (2002). The role of processing limitations in early

identification of specific language impairment. *Topics in language Disorders, 22*(3), 15-29.

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., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur communicative development inventories. *Applied Psycholinguistics*, *21*(1), 95–116.

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

Fisher, C. (2002). Structural limits on verb mapping: the role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science, 5*(1), 55-64. https://doi.org/10.1111/1467-7687.00209

Fisher, C., Hall, D. G., Rakowitz, S., & Gleitman, L. (1994). When it is better to receive than to give: Syntactic and conceptual constraints on vocabulary growth. *Lingua, 92*, 333–375. https://doi.org/10.1016/0024-3841(94)90346-8

Frost, R. L., Monaghan, P., & Christiansen, M. H. (2019). Mark my words: High frequency marker words impact early stages of language learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(10), 1883. https://doi.org/10.1037/xlm0000683

Gentner, D. (1978). On relational meaning: The acquisition of verb meaning. *Child Development*, 988-998.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus

natural partitioning. *Center for the Study of Reading Technical Report; no. 257*.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity and early word learning. In M. Bowerman & S. Levinson (Eds.), *Language Acquisition and Conceptual Development* (pp. 215–256). Cambridge, United Kingdom: Cambridge University Press.

Gentner, D., & Namy, L. L. (2006). Analogical processes in lan- guage learning. *Current Directions in Psychological Science, 15*(6), 297–301. https://doi.org/10.1111/j.1467-8721.2006.00456.x

Gernsbacher, M.A. (1989). Mechanisms that improve referential access. *Cognition, 32*(2), 99-156. https://doi.org/10.1016/0010-0277(89)90001-2

Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition, 1*(1), 3–55. https://doi.org/10.1207/ s15327817la0101_2

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, *17*(3), 311-347. https://doi.org/10.1207/s15516709cog1703_1

Gordon, P. C., & Scearce, K. A. (1995). Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory & Cognition*, *23*(3), 313-323. https://doi.org/10.3758/BF03197233

Hadley, P. A., Rispoli, M., Holt, J. K., Papastratakos, T., Hsu, N., Kubalanza, M., & McKenna, M. M. (2017). Input subject diversity enhances early grammatical growth: Evidence from a parent-implemented intervention. *Language Learning and Development, 13*(1), 54-79. https://doi.org/10.1080/15475441.2016.1193020

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

Hadley, P. & Walsh, K.M. (2014). Toy Talk: Simple strategies to create richer grammatical input. *Language, Speech, and Hearing Services in Schools, 45*(3), 159-172. https://doi.org/10.1044/2014_LSHSS-13-0055

He, A.X., & Arunachalam, S. (2017). Word learning mechanisms. *Wiley Interdisciplinary Reviews: Cognitive Science, 8*(4), e1435. https://doi.org/10.1002/wcs.1435

He, A. X., Kon, M., & Arunachalam, S. (2020). Linguistic context in verb learning: Less is sometimes more. *Language Learning and Development, 16*(1), 22-42. https://doi.org/10.1080/15475441.2019.1676751

He, A. X., & Lidz, J. (2016). When one cue is better than two: Trade-off between processing load and informativity in verb learning and verb extension tasks. Paper Presented at the XIX Biennial International Conference for Infant Studies (ICIS). New Orleans, LA.

Hoff-Ginsberg, E. (1985). Some contributions of mothers' speech to their children's syntactic growth. *Journal of Child Language*, *12*(2), 367-385. https://doi.org/10.1017/S0305000900006486

Horvath, S., & Arunachalam, S. (2019). Optimal contexts for verb learning. *Perspectives of the ASHA Special Interest Groups, 4*(6), 1239-1249. https://doi.org/10.1044/2019_PERSP-19-00088

Horvath, S., Kueser, J.B., Kelly, J., & Borovsky, A. (in revision). Difference or delay? Syntax, semantics, and verb vocabulary development in typically developing and late-talking toddlers. *Language Learning and Development.*

Horvath, S., McDermott, E., Reilly, K., & Arunachalam, S. (2018). Acquisition of verb meaning from syntactic distribution in preschoolers with autism spectrum disorder. *Language,*

Speech, and Hearing Services in Schools, *49*(3S), 668-680. https://doi.org/10.1044/2018_LSHSS-STLT1-17-0126

Horvath, S., Rescorla, L., & Arunachalam, S. (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

Horwitz, S. M., Irwin, J. R., Briggs-Gowan, M. J., Heenan, J. M. B., Mendoza, J., & Carter, A. S. (2003). Language delay in a community cohort of young children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *42*(8), 932-940. https://doi.org/10.1097/01.CHI.0000046889.27264.5E

Hsu, N., Hadley, P. A., & Rispoli, M. (2017). Diversity matters: Parent input predicts toddler verb production. *Journal of child language*, *44*(1), 63. https://doi.org/10.1017/S0305000915000690

Imai, M., Haryu, E., & Okada, H. (2005). Mapping novel nouns and verbs onto dynamic action events: Are verb meanings easier to learn than noun meanings for Japanese children? *Child Development, 76*(2), 340–355. https://doi.org/10.1111/ j.1467-8624.2005.00849_a.x

Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R. M., & Shigematsu, J. (2008). Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child Development*, *79*(4), 979-1000. https://doi.org/10.1111/j.1467-8624.2008.01171.x

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

Jiménez, E., Haebig, E., & Hills, T. T. (2020). Identifying areas of overlap and distinction

797 in early lexical profiles of children with autism spectrum disorder, late talkers, and typical
798 talkers. *Journal of Autism and Developmental Disorders*, 1-17. https://doi.org/10.1007/s10803-
799 020-04772-1
800 Jones, S. S. (2003). Late talkers show no shape bias in a novel name extension task.
801 *Developmental Science*, *6*(5), 477-483. https://doi.org/10.1111/1467-7687.00304
802 Jones, S. S., & Smith, L. B. (2005). Object name learning and object perception: A deficit
803 in late talkers. *Journal of child language*, *32*(1), 223.
804 https://doi.org/10.1017/s0305000904006646
805 Jones, G., & Rowland, C. F. (2017). Diversity not quantity in caregiver speech: Using
806 computational modeling to isolate the effects of the quantity and the diversity of the input on
807 vocabulary growth. *Cognitive Psychology*, *98*, 1-21. 10.1016/j.cogpsych.2017.07.002
808 Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package:
809 Tests in Linear Mixed Effects Models. *Journal of Statistical Software, 82(13),* 1-26.
810 https://doi.org/10.18637/jss.v082.i13
811 Landau, B. & Gleitman, L. (1985). *Language and experience. Evidence from the blind*
812 *child.* Cambridge MA: Harvard University Press.
813 Lenth, R. (2019). emmeans: estimated marginal means, aka least-squares means. R
814 package v. 1.3. 4.
815 Lidz, J., Bunger, A., Leddon, E., Baier, R., & Waxman, S. R. (2009). When one cue is
816 better than two: Lexical vs. syntactic cues to verb learning. *Unpublished manuscript*.
817 Longobardi, E., Spataro, P., Putnick, D. L., & Bornstein, M. H. (2017). Do early noun
818 and verb production predict later verb and noun production? Theoretical implications. *Journal of*
819 *child language*, *44*(2), 480. http://doi.org/10.1017/S0305000916000064

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

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk. Third Edition.* Mahwah, NJ: Lawrence Erlbaum Associates.

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., & Brandone, A. C. (2008). Focusing on the relation: fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science*, *11*(4), 628-634. https://doi.org/10.1111/j.1467-7687.2008.00707.x

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

Mintz, T. H. (2006). Finding the verbs: Distributional cues to categories available to young learners. In K Hirsh-Pasek & R Golinkoff (Eds.), *Action meets word: How children learn verbs* (pp.31-63). Oxford University Press.

Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17*(2), 357–374. https://doi.org/10.1017/s0305000900013817

Naigles, L. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition, 58,* 221–251. https://doi.org/10.1016/0010-0277(95)00681-8

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the society for research in child development*, 1-135. https://doi.org/10.2307/1165788

Nicholas, K., Alt, M., & Hauwiller, E. (2019). Variability of input in preposition learning by preschoolers with developmental language disorder and typically-developing language. *Child Language Teaching and Therapy*, *35*(1), 55-74. https://doi.org/10.1177/0265659019830455

Oakhill, J., & Yuill, N. (1986). Pronoun resolution in skilled and less-skilled comprehenders: Effects of memory load and inferential complexity. *Language and speech*, *29*(1), 25-37. https://doi.org/10.1177/002383098602900104

Naigles, L. R., & Hoff-Ginsberg, E. (1995). Input to verb learning: Evidence for the plausibility of syntactic bootstrapping. *Developmental Psychology*, *31*(5), 827. https://doi.org/10.1037/0012-1649.31.5.827

Naigles, L. R., & Hoff-Ginsberg, E. (1998). Why are some verbs learned before other verbs? Effects of input frequency and structure on children's early verb use. *Journal of child language*, *25*(1), 95-120. https://doi.org/10.1017/s0305000997003358

Oetting, J. B., Rice, M. L., & Swank, L. K. (1995). Quick incidental learning (QUIL) of words by school-age children with and without SLI. *Journal of Speech, Language, and Hearing Research*, *38*(2), 434-445. https://doi.org/10.1044/jshr.3802.434

Onnis, L., Waterfall, H. R., & Edelman, S. (2008). Learn locally, act globally: Learning language from variation set cues. *Cognition, 109*(3), 423-430. https://doi.org/10.1016/j.cognition.2008.10.004

Owen Van Horne, A., & Strother-Garcia, K., (2020). *Distributional properties of the input affect syntax learning for children with Developmental Language Disorder (DLD), but not typically developing (TD) peers.* Poster presentation at the Boston University Conference on Language Development, Boston, MA.

Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn locally,

think globally: Exemplar variability supports higher-order generalization and word learning. *Psychological Science, 21*(12), 1894–1902. https://doi.org/10.1177/ 0956797610389189

Piccin, T., & Waxman, S. (2007). Three-year-olds use cross-situational exposure and contrast to learn verbs. Poster presented at the Society for Research in Child Development Biennial Meeting, Boston, MA.

R Core Team. (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

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., Alley, A., & Christine, J. B. (2001). Word frequencies in toddlers' lexicons. *Journal of Speech, Language, and Hearing Research, 44*, 598–609.

Rice, M., Oetting, J., Marquis, J., Bode, J., & Pae, S. (1994). Frequency of input effects on word comprehension of children with specific language impairment. *Journal of Speech and Hearing Research, 37*, 106-122. https://doi.org/10.1044/jshr.3701.106

Riches, N. G., Tomasello, M., & Conti-Ramsden, G. (2005). Verb learning in children with SLI: Frequency and spacing effects. Journal of Speech, Language, and Hearing Research, 48(6), 1397-1411. https://doi.org/10.1044/1092-4388(2005/097)

Robins, D. L., Fein, D., & Barton, M. (2009). Modified checklist for autism in toddlers, revised, with follow-up (M-CHAT-R/F). Self-published.

Scheffner Hammer, C., Morgan, P., Farkas, G., Hillemeier, M., Bitetti, D., & Maczugab, S. (2017). Late talkers: A Population-based study of risk factors and school readiness consequences. *Journal of Speech, Language, and Hearing Research, 60*(3), 607-626.

https://doi.org/10.1044/2016_JSLHR-L-15-0417

Scott, R. M., & Fisher, C. (2009). Two-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language and Cognitive Processes, 24*(6), 777–803. https://doi.org/10.1080/ 01690960802573236

Shi, R., & Melançon, A. (2010). Syntactic categorization in French-learning infants. *Infancy*, *15*(5), 517-533. https://doi.org/10.1111/j.1532-7078.2009.00022.x

Sims, C. E., Schilling, S. M., & Colunga, E. (2017). Dynamics of Word Learning in Typically Developing, Temporarily Delayed, and Persistent Late-Talking Children. In *Neurocomputational Models of Cognitive Development and Processing: Proceedings of the 14th Neural Computation and Psychology Workshop* (pp. 83-97).

Suppes, P. (1974). The semantics of children's language. *American Psychologist, 29*, 103–114.

Syrett, K., Arunachalam, S., & Waxman, S. R. (2014). Slowly but surely: Adverbs support verb learning in 2-year-olds. *Language Learning and Development, 10*(3), 263–278. https://doi.org/10.1080/15475441.2013.840493

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, 1454-1463. https://doi.org/10.2307/1130423.

Twomey, K. E., Lush, L., Pearce, R., & Horst, J. S. (2014). Visual variability affects early verb learning. *British Journal of Developmental Psychology, 32*(3), 359–366. https://doi.org/10.1111/ bjdp.12042

Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology, 36*(5), 571–581. https://doi.org/10.1037//0012-1649.36.5.571

912 Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old
913 infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology,*
914 *59*(1), 67–95. https://doi.org/10.1016/j.cogpsych.2009.02.001
915 Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool language scales* (5th
916 ed.). San Antonio, TX: Pearson/PsychCorp.
917 Zubrick, S.R., Catherine L. Taylor, L.C., Rice, M., & Slegers, D. (2007). Late language
918 emergence at 24 months: An epidemiological study of prevalence, predictors, and covariates.
919 *Journal of Speech, Language, and Hearing Research, 50*(6), 1562-1592.
920 https://doi.org/10.1044/1092-4388(2007/106)
921

922 Table 1. Participating children's scores on standard assessments.

	MCDI Level 2 Words Produced		PLS-AC Standard Score		PLS-EC Standard Score	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*
All children (N = 47)	68	28	107	16	107	17
Late Talkers (n = 14)	38	26	91	14	91	11
Typically Developing (n = 33)	88	17	114	12	114	12

923

924

925

926 Table 2. Descriptions of the visual stimuli on each experimental trial. One trial (unlisted) was

927 removed due to visual stimuli design error.

Novel Verb	Familiarization Scene	Preserved Action	Preserved Object
nork	A man sliding a block in a circle on a table	A man sliding a small cup in a circle on a table	A man hopping a block across a table
pell	A boy rocking a bunny backward and forward	A boy rocking a small teddy bear backward and forward	A boy spinning a bunny in a circle
pilk	A girl pushing a small chair across the floor	A girl pushing a box across the floor	A girl tipping a chair up and down
sem	A boy tossing an apple up and down	A boy tossing a ball up and down	A boy sliding an apple back and forth on a table
tope	A lady flipping over a large cup	A lady flipping over a box	A lady pushing a large cup forward
wug	A girl hitting a large flower with her hand	A girl hitting a ball with her hand	A girl waving a flower from side to side
ziff	A girl raising and lowering a dump truck	A girl raising and lowering a large teddy bear	A girl tipping the bed of a dump truck

928

929 Table 3. Parameter estimates for linear mixed-effects model with vocabulary treated as a
930 continuous variable. Note that the values for the parameter estimates are very small because of
931 the scale of the raw vocabulary score, which ranges from 0-100.

	Estimate	SE	*t*-value	*p*-value
Intercept	-0.28	0.705	-0.396	
Time, sec	-0.063	0.400	-0.16	.87
Condition	1.95	0.395	4.94	<.0001*
Vocabulary	0.00165	0.0081	0.203	.84
Time x Condition	-1.00	0.222	-4.51	<.0001*
Time x Vocabulary	-0.00129	0.00451	-0.285	.77
Condition x Vocabulary	-0.0218	0.0054	-4.047	<.0001*
Time x Condition x Vocabulary	0.0126	0.00304	4.16	<.0001*
Age (months, centered around mean)	-0.0612	0.028	-2.203	.028*

932

933

934 Table 4. Parameter estimates for linear mixed-effects model with children grouped as either late
935 talkers or typically developing.

	Estimate	SE	*t*-value	*p*-value
Intercept	-0.42	0.55	-0.76	
Time, sec	-0.04	0.31	-0.13	.89
Condition	1.51	0.26	5.72	<.0001*
Group (TD)	0.36	0.48	0.76	.44
Time x Condition	-0.72	0.15	-4.88	<.0001*
Time x Group (TD)	-0.16	0.26	-0.61	.54
Condition x Group (TD)	-1.52	0.32	-4.77	<.0001*
Time x Condition x Group (TD)	0.84	0.18	4.72	<.0001*

936

937 Figure 1. A schematic depiction of one trial (of eight).

938

939 Figure 2. Children's gaze to the Familiar Action scene during the Test phase by condition. Error
940 bars indicate standard error of participant means.
941

942 Figure 3A. Late talkers' gaze to the Familiar Action scene during the Test phase by condition.

943 Error bars indicate standard error of participant means.

944

945

946 Figure 3B. Typically developing children's gaze to the Familiar Action scene during the Test
947 phase by condition. Error bars indicate standard error of participant means.
948