Linguistic Context in Verb Learning: Less is Sometimes More

Angela Xiaoxue He, Maxwell Kon & Sudha Arunachalam

Routledge
Taylor & Francis Group

Linguistic Context in Verb Learning: Less is Sometimes More

Angela Xiaoxue He[a], Maxwell Kon[b], and Sudha Arunachalam[c]

[a]School of Philosophy and Department of Linguistics, University of Southern California, Los Angeles, California, USA; [b]Department of Speech, Language, and Hearing Sciences, Boston University, Boston, Massachusetts, USA; [c]Communicative Sciences of Disorders, New York University

ABSTRACT

Linguistic contexts provide useful information about verb meanings by narrowing the space of candidate concepts. Intuitively, the more information, the better. For example, "the tall girl is *fezz*ing," as compared to "the girl is *fezz*ing," provides more information about which event, out of multiple candidate events, is being labeled; thus, we may expect it to better facilitate verb learning. However, we find evidence to the contrary: in a verb learning study, preschoolers (N = 60, mean age = 38 months) only performed above chance when the subject was an unmodified determiner phase, but not when it was modified (Experiment 1). Experiment 2 replicated this pattern with a different set of stimuli and a wider age range (N = 60, mean age = 45 months). Further, in Experiment 2, we looked at both learning outcomes—by evaluating pointing responses at Test, and also the learning process—by tracking eye gaze during Familiarization. The results suggest that children's limited processing abilities are to blame for poor learning outcomes, but that a nuanced understanding of how processing affects learning is required.

Introduction

Identifying the meaning of a new word is not a straightforward task, because the context of language use makes available infinitely many possible meanings (Chomsky, 1959; van Quine, 1960). For nouns that label common objects, this problem may be mitigated by shared human abilities such as attention to social cues or biases that constrain hypotheses about word meaning (e.g., Clark, 1987; Markman, 1991; Markman & Hutchinson, 1984). But these resources may be markedly less useful for learning verb meanings. In a task where adults were asked to "simulate" child language acquisition by observing muted videos of parent-child interactions and guessing what words the parents were uttering, Gillette, Gleitman, Gleitman, and Lederer (1999) found that adults correctly identified nouns three times more often than verbs, identifying verbs only 7.7% of the time. They inferred that determining the target referent of a verb by observing the world alone is profoundly difficult.

The verb's linguistic context, however, can play an important role in alleviating this difficulty. Gillette et al. (1999) also found that adults given linguistic information along with the muted videos did much better. A list of the verb's co-occurring nouns (e.g., "Daddy," "cookie") increased correct identification to 29% of the time, the verb's syntactic frame with nonsense morphemes replacing the content morphemes (e.g., *Ver* gonna __ the *telfa*) increased it to 51.7%, and both cues together (e.g., Daddy's gonna __ the cookie) increased it to 75%. Gleitman and her colleagues argue that linguistic context serves as a "zoom lens" to help learners narrow the search space of potential event referents for an unfamiliar verb (Gleitman, 1990; Landau & Gleitman, 1985). Structurally, the verb's argument structure offers a cue to the broad type of event it describes; for example, "it's *gorp*ing it" describes a

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/hlld.

2-participant event. Semantically, the lexical content of the arguments reveals potential referents of those participants; for example, "the duck is *gorp*ing the bunny" describes an event involving those two participants, and "she is *gorp*ing him," while less specific about their identities, cues their genders. The zoom-lens effect will be particularly helpful when the observational situation is cluttered. For example, on a crowded playground with multiple people engaged in different actions, a richly informative linguistic context (e.g., "The girl in the red jacket is swinging") may allow the listener to identify a specific agent, and thus, the event being labeled.

Therefore, for successful verb learning, the information contained in linguistic context is essential, not only for adults, but also for children (see Piccin and Waxman (2007) for a replication of Gillette et al.'s (1999) task with 7-year-olds). Decades of research have shown that young learners can use linguistic context in the service of verb learning, both structural information such as the number of arguments (e.g., Fisher, 1996, 2002; Fisher, Hall, Rakowitz, & Gleitman, 1994; Hirsh-Pasek & Golinkoff, 1996; Naigles, 1990; Yuan, Fisher, & Snedeker, 2012) and semantic information such as lexical content (e.g., Arunachalam & Waxman, 2011, 2015; Fisher et al., 1994; Imai, Haryu, & Okada, 2005; Imai et al., 2008; Syrett, Arunachalam, & Waxman, 2014).

But what kind of linguistic context is most supportive for verb learning in young children? With respect to semantic information, we might intuitively think that the more information the linguistic context carries (as long as that information is truthful), the more focused the zoom lens will be. Some findings are consistent with this intuition. For Arunachalam and Waxman (2011, 2015)) taught English-acquiring 2-year-olds novel transitive verbs describing events in which an actor acted on an object. At Test, two new events were presented simultaneously, one depicting the same actor and action but a new object, and the other depicting the same actor and object but a new action. They found that children learned verbs appearing in semantically richer contexts—sentences with lexical determiner phrases (DPs) as subjects and objects (e. g., "The boy is *pilk*ing a balloon") – but failed when given semantically sparser contexts with pronominal arguments (e.g., "He is *pilk*ing it"). Similarly, Imai et al. (2008) found that older children, 5-year-old English learners, could succeed with pronominal arguments, but not with null arguments (e.g., "*Pilk*ing"), again indicating that more information was better than less. These studies, taken together, document that young children require substantial semantic support from linguistic context to discover the meaning of a novel verb, although the specific amount of semantic information they need may decrease with development.

However, other findings suggest that more information may not always be better. In a replication of Arunachalam and Waxman (2011) with 24-month-old Korean learners, the opposite pattern obtained from the English results. Korean learners performed better when the novel transitive verb occurred with its arguments elided (e.g., "*Pilk*ing") than when both arguments were overt, though performance was not above chance in either condition (Arunachalam, Leddon, Song, Lee, & Waxman, 2013). Again, Imai and colleagues found a similar pattern with older children: Japanese-learning 5-year-olds succeeded with elided arguments but not with pronouns (Imai et al., 2005, 2008).

Similar evidence also comes from intransitive verbs. Lidz, Bunger, Leddon, Baier, and Waxman (2009) showed that younger English-learning children (22 months) learned novel intransitive verbs with a pronominal subject (e.g., "It's *blick*ing") but not with a lexical DP subject (e.g., "The flower is *blick*ing"). He and Lidz (2016) replicated this finding and further showed that the difference was due to semantic content rather than syntactic complexity, because children also successfully learned the verbs from sentences like, "That thing is *blick*ing," where the subject was semantically lighter but structurally complex.

These findings all suggest a seemingly counter-intuitive conclusion—–that sometimes less information is better. But why is that so? We suggest that informativity, which we define as the amount of (truthful) information in the utterance, interacts with other factors and that the balance among these factors determines how supportive a linguistic context may be for verb learning.[1] Specifically, we

---

[1]Our use of the term "informativity" differs from that in the pragmatics literature (i.e. the Quantity Maxim, Grice, 1975). According to the Quantity Maxim, the speaker provides as much information as is needed, no more (otherwise "over-informative") and no less (otherwise "under-informative"). We use "informativity" in a similar sense as the term "information density" in the computational linguistics literature, which describes the amount of information conveyed in an utterance (Aylett & Turk,

must also consider how informativity interacts with factors that determine how *useful* (i.e., how relevant for identifying the verb's referent), or how *usable* (i.e., how interpretable or decodable), that information is.

With respect to *usefulness*, one relevant factor may be pragmatic felicity of the sentence containing the novel verb. The verb learning studies reviewed above presented novel verbs in visual contexts with only one event (e.g., a boy acting on a balloon, a flower moving) and thus there was only one salient candidate referent for the novel verb, and only minimal disambiguating information was needed in the linguistic context. Extra information may have rendered the sentence "overinformative" and therefore pragmatically infelicitous. This would explain why learning intransitive verbs may require less linguistic information than transitive verbs: to acquire transitive verbs, learners must identify who is doing what to whom, requiring some specificity in the linguistic context, which may be overinformative for learning intransitive verbs whose event structure is much simpler. It may also explain why learners of argument-drop languages may do better with less information—overtly mentioning arguments that are often omitted in their language may be pragmatically disfavored.

However, there is no direct evidence that child learners are sensitive to this kind of pragmatic infelicity. Child-directed speech is replete with redundancy (e.g., Newport, 1975; Snow, 1972), and in experimental studies, children up to 5 years of age are tolerant of overinformative statements (Davies & Katsos, 2010; Morisseau, Davies, & Matthews, 2013). For example, Morisseau et al. (2013) found that when provided an overinformative instruction like "Find the bird with the feather" when there was only one normal-looking bird in the display, 3-year-olds did not appear to notice the infelicity and were not delayed in their response.

With respect to *usability* of information, young learners may be sensitive to the processing load associated with the linguistic context. To benefit from the information in the linguistic context, children must be able to process it, but young children are still honing their processing skills (e.g., Trueswell, Sekerina, Hill, & Logrip, 1999); they must parse to learn while they are still learning to parse (Kidd, Bavin, & Brandt, 2013; Omaki & Lidz, 2015; Trueswell & Gleitman, 2007). Could more information hurt verb learning by virtue of incurring a heavier processing load? In the case of the intransitive verbs studied by Lidz et al. (2009) and He and Lidz (2016), the lexical DP may have incurred too great a processing load for these quite young learners, while the pronominal subject was sufficiently informative (given that there was only one entity in the scene) and easier to process. In support of this interpretation, these authors found that children performed better with the lexical DP if they were exposed to it beforehand (e.g., "Look at the flower … . The flower is *blick*ing"). They surmised that this pre-exposure lessened the processing burden associated with retrieving the lexical DP's referent. In the case of verb learning in Korean (Arunachalam et al., 2013) and Japanese (Imai et al., 2005, 2008) reviewed above, the availability of argument drop in these languages may mean that learners have less experience, and therefore more difficulty, processing overt arguments (Arunachalam & Waxman, 2011, 2015).

Independent evidence from language comprehension studies outside the domain of verb learning supports the hypothesis that limitations on the developing parser affect children's use of linguistic context. Choi and Trueswell (2010) demonstrated that 4- and 5-year-olds had difficulty inhibiting misinterpretations of garden-path sentences despite highly informative cues from the sentence, attributable to their limited processing abilities. Huang and Arnold (2016) found decreased sensitivity to syntactic cues in 5-year-olds when the sentence imposed a particularly difficult parsing challenge (i.e., passive sentences that require revising an initial thematic role assignment). More relevantly to the current study, Fernald, Marchman, and Hurtado (2008) found that children who were faster to process through a modified DP were more likely to learn the meaning of a novel noun downstream in the sentence (as in, "There's a blue cup on the *deebo*," given a scene with a blue cup on one novel object and a red cup on another). Thus, the ability to process familiar words in a

---

2004; Jaeger & Levy, 2007). Thus, an utterance that contains more truthful information is by our definition higher in informativity.

linguistic context facilitates acquisition of new words encountered later. In the case of verbs, whose referents are often dynamic events, more efficient processing of familiar words early in the sentence may be even more important, allowing the child to shift attention to the correct event in time to see it unfold.

Therefore, informativity, pragmatic felicity, and processability may all play a role in determining whether a particular linguistic context will be supportive for verb learning. While young children usually need significant semantic support to discover the meanings of novel verbs, it may be that "less information is more" when the linguistic information is either not useful or not usable by the young learner.

In the current study, we explore these issues using an experimental verb learning task similar to those used in the studies reviewed above. However, in these prior studies, the visual context depicting the verb was so simple (i.e., one event with one salient action) that there were very few candidate referents. The linguistic zoom lens is likely to play a severely diminished role in such contexts. For an adult, at least, the barest of linguistic contexts (e.g., "*Pilk*ing") would have been sufficient. Given that in the real world, the extralinguistic context is likely to be much more complex than that in a laboratory setting (Gillette et al., 1999; Gleitman, 1990; Landau & Gleitman, 1985), we take a first step toward a more complex learning scenario by introducing novel verbs with two simultaneous events (e.g., girl waving, boy clapping), only one of which is labeled by the verb. Thus, the linguistic context is crucial for identifying the referent event. We ask: when learners encounter novel verbs in a setting where they *must* use the linguistic context to identify the verb's referent event, is more information better? Or is there still a tradeoff between informativity and pragmatic felicity and/or processing load? A second goal of the current study is to tease apart pragmatics and processing. Above, we interpreted examples of children's relatively poorer performance with more information to be either about pragmatics or about processing. Here, we include a manipulation to test both possibilities.

Current study

In the current study, we study young children's acquisition of novel intransitive verbs in linguistic contexts that vary in informativity, pragmatic felicity, and processability. We focus on intransitive verbs both because their simple argument structure allows us to manipulate only one argument, and because their simple event structure means that linguistic context can be used primarily to identify which referent event is being labeled rather than how multiple event participants relate to each other.

We vary the verbs' linguistic contexts by manipulating the semantic content contained in the subject DP. As reviewed above, one salient way in which linguistic contexts can differ in the amount of semantic content is how the verb's arguments are realized (e.g., as pronouns, content nouns, or null arguments). In the current study, we present subject arguments with contentful DPs either with or without an adjectival modifier, as in (a) and (b) below. Sentence (b) contains more semantic content—that is, a more detailed description of the event participant—than sentence (a) and is thus higher in informativity. (Recall that we define "informativity" as the amount of semantic information contained, regardless of whether that information is useful for identifying the target.)

(a) The girl is *fezz*ing.
(b) The tall girl is *fezz*ing.

The novel verb is presented alongside two visual scenes that play simultaneously on either side of the screen, requiring learners to attend to the linguistic context to learn which is being referred to, and in turn to learn the verb's meaning. We manipulate the visual context in order to study the potential effects of pragmatic felicity. In one condition, the agents in the two scenes come from different basic-level categories but have contrasting salient properties (e.g., a tall girl in one scene and a short boy in another), such that the two scenes are distinguishable by the noun alone in the DP while the modifier (if any) also contributes true

information. In another, the agents also have contrasting salient properties but come from the same basic-level category (e.g. a tall girl in one scene and a short girl in another), such that the two scenes are distinguishable only by the modifier. This allows us to see how verb learning is affected by overly informative linguistic contexts (when the modifier is unnecessary) as compared to appropriately informative contexts (when the modifier is necessary).

Given these two types of linguistic context and two types of visual context, we designed three verb learning conditions,[2] varying only in the Familiarization phase. In the first condition, which we call the Light condition (i.e., semantically lighter), children hear novel verbs in sentences like (a), with visual scenes with different basic-level-category agents. In the second condition, which we call the Heavy-Unnecessary condition (i.e., semantically heavier, and the modifier is not necessary for identifying the target event), children hear sentences like (b) but see scenes with different basic-level-category agents. In the third condition, which we call the Heavy-Necessary condition (i.e., semantically heavier, and the modifier is necessary for identifying the target event), children hear the verbs in sentences like (b) but see scenes with agents from the same basic-level category. (The fourth logical possibility given our visual and linguistic manipulations is a "light" linguistic stimulus with two agents from the same basic-level category, but we do not include it because it would be unresolvable.) The Test phase in all conditions is identical: children see two new scenes, each depicting a new agent performing one of the actions seen during Familiarization, and they are asked, for example, to "point to *fezzing*."

In all three conditions, because there are two candidate events when the verb is first introduced, the linguistic context is needed for identifying the verb's referent. If it is the case that more information is better, we would expect better verb learning in the two Heavy conditions than in the Light condition. However, the two Heavy conditions differ in whether the modifier is necessary for identifying the verb's referent. If children are sensitive to pragmatic felicity, we would expect better performance in the Heavy-Necessary condition than the Heavy-Unnecessary condition. Alternatively, because the two Heavy conditions, by virtue of having "heavier" subject DPs, impose a higher processing load as compared to the Light condition, children may only succeed in the Light condition and not in the two Heavy conditions if processing load limits their use of the linguistic context.

We report two experiments. In Experiment 1, like prior verb learning work, we looked at children's verb learning *outcomes*—that is, their choice of the target versus the distractor scene during the Test phase. Results suggested an intriguing trend that more information might hinder, rather than facilitate, verb learning, and that processing load, rather than pragmatics, might be accountable. To replicate and to better understand the results in Experiment 1, in Experiment 2, we looked not only at learning outcomes but also at the learning *process*—by tracking eye gaze during the Familiarization phase to make inferences about which part of the learning task they struggled with. Data and analysis codes for both experiments can be found at osf.io/tv7kh.

Experiment 1

Participants

Sixty English-learning children (30 boys, 30 girls) with a mean age of 38.2 months (range: 30.4–48.0 months) were included in the final sample. They were all reported by their parent to be exposed to English at least 70% of the time, and to have no known language, communication, or uncorrected hearing or vision problems. Data from 13 additional children were excluded due to failure to point correctly on at least one of two pointing training trials, or failure to point at all on at least 2 of the 4 experimental trials.

---

[2]"Light" and "Heavy" in the condition names distinguish conditions that vary in informativity: "Light" is lower and "Heavy" is higher in informativity. "Necessary" and "unnecessary" distinguish conditions that vary in pragmatic felicity: "necessary" follows the Quantity Maxim (Grice, 1975) but "unnecessary" provides over-informative statements.

Stimuli

Stimuli consisted of two pointing training trials and four experimental trials. The visual stimuli for training trials were video clips of familiar characters (e.g., Big Bird), and auditory prompts were provided by an experimenter. For experimental trials, the videos were recordings of actors engaging in simple actions (e.g., a girl waving) or objects manipulated by a hand (e.g., a truck rotating).[3] In the latter case, the hand was backgrounded so that the object, not the agent, was the only salient event participant. Auditory stimuli were recorded by a female native English speaker using child-directed speech, edited in Praat (Boersma & Weeknink, 2014), and combined with the visual stimuli in Final Cut Pro X.

Apparatus and procedure

The child and parent were first welcomed into a playroom where the child played with an experimenter and the parent provided consent and completed the MacArthur Communicative Development Inventory Level III (MCDI) together with a supplementary vocabulary checklist including the adjectival modifiers used in the experiment. MCDI scores were not significantly different across conditions; see Table 1. Then, in the testing room, the child sat in a car seat or on the parent's lap 18 inches from a 24-inch monitor. Stimuli were presented from a desktop PC (Dell Precision T5500). The parent was asked not to talk, and to wear a blindfold if the child was on the parent's lap. An experimenter sat next to the child to elicit and record pointing responses; another experimenter also recorded pointing from behind a curtain via integrated webcam feed.

Pointing training

The pointing training session introduced familiar characters engaged in familiar actions. Children saw two video clips side-by-side and were asked to point to one (e.g., "Show me dancing.").

Experimental trials

For experimental trials, children were randomly assigned to one of three conditions in a between-subject design: Light, Heavy-Unnecessary, or Heavy-Necessary (each N = 20). They participated in four trials, each comprising a Familiarization phase and a Test phase. Trials differed by condition only in the Familiarization phase. All children completed four trials in the same order. See Table 2 for an example trial and Table 3 for a summary of stimuli in all trials.
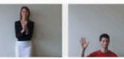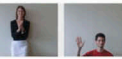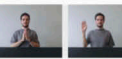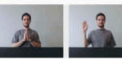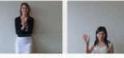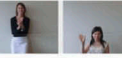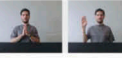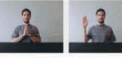
Familiarization phase. On each trial, children first saw two side-by-side video clips, each depicting a different event participant and a different action (e.g., a tall girl waving). Children heard a novel verb in a sentence describing one of the video clips. The sentence either had an unmodified subject DP (Light condition, e.g., "The girl is *fezz*ing,") or one with a modifier (Heavy conditions, e.g., "The

Table 1. Summary of MCDI-III vocabulary scores of children in Experiment 1.

	Light (n = 20)		Heavy-Unnecessary (n = 19)[4]		Heavy-Necessary (n = 20)	
	Mean	SD	Mean	SD	Mean	SD
Total vocabulary	74.05	12.60	79.32	17.26	74.45	21.13
Nouns	39.20	6.50	41.68	6.91	38.85	9.48
Verbs	9.90	1.71	10.00	2.75	8.95	3.43

[3]We did not use wholly novel actions because it is difficult to come up with novel actions that could be plausibly lexicalized by a single verb in a human language, and because there is no evidence that children engage in mutual exclusivity when learning novel verbs as they do for nouns (Hesketh & Ellis Weismer, 1997; Merriman, Marazita, & Jarvis, 1995).

Table 2. An example trial (Experiment 1).

		Familiarization			Test			
		6 sec	.5 sec	6 sec	.5 sec	6 sec	1 sec	6 sec
Visual stimuli	Light & Heavy-Unnecessary conditions						⭐	
	Heavy-Necessary condition						⭐	
Auditory stimuli	Light condition	The girl is *fezz*ing.						
	Heavy-Unnecessary & Heavy-Necessary conditions	The tall girl is *fezz*ing.	(silence)	(silence)	Take a look!	(silence)	Show me *fezz*ing!	Point to *fezz*ing!

Table 3. A summary of stimuli used in all trials in Experiment 1.

	Actors			
Trial	Light & Heavy-Unnecessary conditions	Heavy-Necessary conditions	Actions	Sentence
1	tall girl, short boy	tall girl, short girl	clap, wave	"The (tall) girl is *fezz*ing"
2	happy boy, sad girl	happy boy, sad boy	march, crouch	"The (happy) boy is *moop*ing"
3	round ball, rectangular dump truck	round ball, football	tilt up/down, roll back/forth	"The (round) ball is *leam*ing"
4	big baby doll, little pig	big baby doll, little baby doll	lean forward, bounce	"The (big) baby is *gopp*ing"

tall girl is *fezz*ing"). We chose adjectives that we expected children to know, avoiding color terms because they are notoriously slow to be acquired (e.g., Backscheider & Shatz, 1993; Wagner, Dobkins, & Barner, 2012). Parental report confirmed that these adjectives were in most participants' productive vocabularies; see Table 4.[5] The videos played for 6 seconds along with the sentence, and then again for 6 seconds in silence after a brief blank screen. Children only heard the sentence with the novel verb once, making this quite a demanding task.

Test phase. Two new scenes were then presented, featuring a different actor from the one seen previously, but performing the same two actions. After the attention-getting audio, "Take a look!", the child was prompted to point to the scene labeled by the verb, twice ("Show me *fezz*ing" and "Point to *fezz*ing"). If the child did not point after the two prerecorded prompts, a third prompt was offered by the experimenter seated next to the child. The scene displaying the same action as the one labeled by the novel verb during Familiarization was the *target*, and the other was the *distractor*. The target scene for each trial was the same

Table 4. Percentage of participants reported to produce the adjectives used in Experiment 1.

	Light (n = 20)	Heavy-Unnecessary (n = 19)[6]	Heavy-Necessary (n = 20)
tall	90	90	85
happy	95	90	95
round	na	na	na
big	100	95	100

[4]One parent of a participant in the Light condition did not provide vocabulary data.
[5]One of the eight adjectives, "round," was inadvertently omitted from the vocabulary checklist.

for all children (e.g., for all children, the *clapping* action was the target on the *fez* trial). The left-right position of the target and distractor at Test, as well as whether it appeared on the same or opposite side as it had during Familiarization, was counterbalanced across trials.

Coding

Both experimenters independently recorded whether the child pointed to the left, right or not at all. Disagreements were resolved by discussion immediately after the session (specifically, the coder behind the curtain sometimes coded left and right from his/her own perspective, which was the mirror image of the child's; this systematic discrepancy was easily corrected). The first point (if more than one) produced on each trial served as the dependent measure: 0 for pointing to the distractor, and 1 for the target. The 7.5% (18 out of 240) of trials on which children did not point were excluded from analysis.

Results

The results are illustrated in Figure 1. Mean pointing score in the Light condition was 0.65 ($SD = 0.29$), in the Heavy-Unnecessary condition, 0.53 ($SD = 0.30$), and in the Heavy-Necessary condition, 0.52 ($SD = 0.26$). In the Light condition, 18 children chose the target on at least half of the trials, among which 6 did so on all four trials. In the Heavy-Unnecessary condition, only 12 chose the target on at least half of the trials, and 3 did so on all four. In the Heavy-Necessary condition, 15 chose the target on at least half of the trials, and 2 did so on all four. For all statistical analyses, we adopted a significance level of 0.05.[7] Analyses were carried out in R (version 3.3.3) using the lme4 package (version 1.1–12).

To determine whether children learned the novel verbs in any condition, we compared their performance to chance level (i.e., 0.5). Data from each condition were submitted separately to
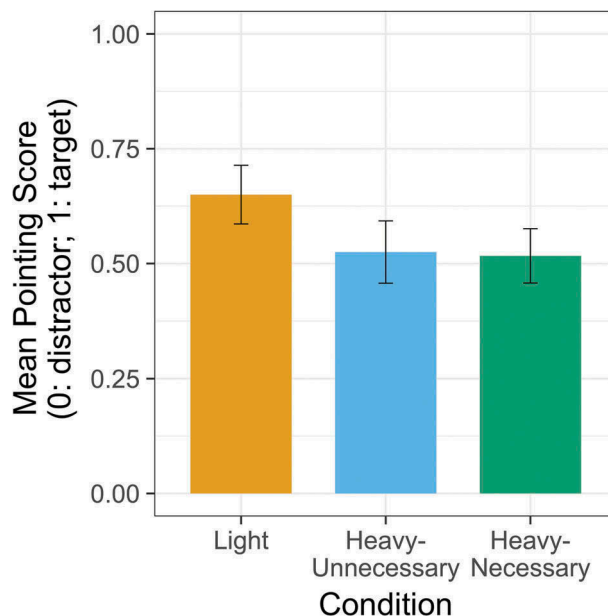
Figure 1. Mean pointing score in Experiment 1 (Error bars indicate standard errors of participant means).

---

[6]One parent of a participant in the Light condition did not provide vocabulary data.
[7]We also reported effects with a *p*-value between 0.05 and 0.1 but did not use them as bases for interpretation.

Table 5. Summary of by-trial performance in Experiment 1.

	Light		Heavy-Unnecessary		Heavy-Necessary	
	Mean	SD	Mean	SD	Mean	SD
Trial 1	0.47	0.51	0.42	0.51	0.42	0.51
Trial 2	0.63	0.50	0.47	0.51	0.47	0.51
Trial 3	0.80	0.41	0.47	0.51	0.50	0.52
Trial 4	0.70	0.47	0.79	0.42	0.72	0.46

mixed-effects logistic regression models, which allowed us to add participant and trial as random factors and age (in months; centered around its mean) as a continuous predictor. We compared pointing responses to chance by evaluating the intercept parameter. Only the Light condition yielded above-chance learning (intercept parameter = 0.74, $z = 2.03$, $p < .05$); the Heavy-Unnecessary condition (intercept parameter = 0.17, $z = 0.45$, $p = .65$), and the Heavy-Necessary condition (intercept parameter = 0.11, $z = 0.47$, $p = .64$) did not. There was no effect of age in any condition. Correlation analyses were conducted to see whether there was any correlation between vocabulary (total MCDI vocabulary, as well as number of modifiers known) and performance. However, we found no such correlation—all p-values for Pearson's R correlation were greater than 0.23, with one exception—in the Heavy-Necessary condition, the correlation between MCDI vocabulary and performance approached significance ($R^2 = 0.19$, $p = .055$).

Next, we asked whether children improved across trials by adding block (first two trials vs. last two trials) as a fixed effect to the models. We found no main effect of block in either the Heavy-Unnecessary condition ($z = 1.57$, $p = .12$) or the Heavy-Necessary condition ($z = 1.42$, $p = .15$), but a marginally significant effect ($z = 1.91$, $p = .056$) in the Light condition. A closer look at each trial individually, however, does seem to reveal some across-trial improvement. In particular, children in all conditions performed at chance level on Trial 1, and above chance on Trial 4; on Trials 2 and 3, a similar pattern as the overall pattern obtained—children showed above-chance performance in the Light condition but not in either of the two Heavy conditions. In other words, although there were no significant effects of block, the trends suggest that children may have improved over the course of the study, but more quickly in the Light condition than in the Heavy conditions. See Table 5 for a summary of performance by trial.

Finally, we entered the data from all conditions into a single mixed-effect logistic regression model to look for significant differences among conditions, but we found none: there was no significant difference between the Light and Heavy-Unnecessary conditions ($z = -1.29$, $p = .20$), nor between the Light and Heavy-Necessary conditions ($z = -1.42$, $p = .15$), nor between the two Heavy conditions ($z = 0.13$, $p = .90$).

Discussion

Taken together, the analyses that compared performance in each condition to chance level suggest that children learned novel verbs introduced with unmodified DPs, but not with modified DPs. This suggests that more information may not be better, even when information in the linguistic context is helpful for determining the meaning of the novel verb. However, the lack of a significant difference between conditions prevents us from drawing strong conclusions. This null result may be due to the small range of values possible in a binary pointing task, and the relatively low "ceiling" with young children.[8] Nevertheless, replication of these results is needed, which we take up in Experiment 2.

We considered two possible explanations for children's difficulty with the modified conditions. One is pragmatics—perhaps children found the modified DP pragmatically infelicitous. We tested this possibility by including two conditions, each with a modified DP, but one with a visual scene

---

[8]The range of performance was 53%-65%, similarly to other verb learning studies using these tasks; Arunachalam and Waxman (2015) and Syrett et al. (2014) each had a 65% success rate in the condition with the best performance.

that made the modifier necessary (Heavy-Necessary condition) and another with a visual scene that made the modifier unnecessary and therefore infelicitous (Heavy-Unnecessary condition). Children struggled equally in both, indicating that pragmatics is not a likely explanation for their difficulty with the heavy conditions. Instead, the results are consistent with an account taking into consideration the processing load that information imposes on children's developing parsers. Specifically, the extra information carried by the modified DPs may have overloaded children's parsers, preventing them from learning the novel verbs downstream.

There are at least two ways in which modified DPs can hinder parsing. One possibility is that children simply fail to parse through the modified DPs. This seems unlikely because Thorpe and Fernald (2006) demonstrated that even 24-month-olds are able to process through DPs with one modifier. A second possibility, then, is that children can parse the modified DP, but are left with insufficient resources to learn the novel verb. Fernald et al. (2008) found in a noun learning study that 36-month-olds were able to process modified DPs and use that information to learn novel nouns downstream. However, prior work on verb learning has established that it is typically even more demanding than noun learning (e.g., Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005), for reasons that apply both in general as well as specifically in our experimental manipulations. First, verbs' referents are often dynamic events and were depicted with dynamic video scenes in our study, while the referents in Fernald et al. (2008) were novel objects depicted with static images. Dynamic events require children to process the visual scene temporally, and may suggest many possible interpretations as the event unfolds. Second, verb acquisition requires encoding argument structure and event structure information, while the referents of object nouns (like those tested in Fernald et al., 2008) are inherently less complex. Finally, learning any kind of word requires generalizing from a single exemplar to other members of the same category, but there is substantial evidence that children find this more difficult for verbs that label dynamic events than for nouns that label objects (e.g., Behrend, 1990; Forbes & Farrar, 1993, 1995; Imai et al., 2005, 2008; Kersten & Smith, 2002). While in Fernald et al.'s (2008) study, children saw the same novel objects depicted from familiarization to test, our design required children to generalize the new verb to a new agent. Therefore, we suspect that the difficulty of verb learning, both in general and in the specific task children faced in our study, is responsible for children's difficulty in the Heavy conditions. They lacked sufficient processing resources after processing the modified DPs to permit them to tackle the challenging verb learning task. In Experiment 2, we sought to replicate Experiment 1, and also to pursue this hypothesis about the reason for children's failure to learn verbs with modified DP subjects.

Experiment 2

To replicate the findings of Experiment 1, we used the same design, but with a new set of stimuli and a slightly broader age range. To obtain insight into why children struggled with the Heavy conditions in Experiment 1, we asked whether, as in Thorpe and Fernald (2006) and Fernald et al. (2008), children were able to process the subject DPs. If so, we have indirect evidence that the barrier to verb learning was not identifying the referent of the modified DPs, but rather the accumulated processing load that doing so encumbered. To accomplish this, we looked not only at children's pointing responses at Test as an index of their verb learning outcomes, but also their eye gaze during Familiarization as an index of their learning process.

Participants

Sixty English-learning children (30 boys, 30 girls) with a mean age of 45.2 months (range: 30.5–59.1 months) who did not participate in Experiment 1 were included in the final sample. We included a slightly broader age range than Experiment 1 to look for evidence of developmental change. As in Experiment 1, children were reported by their parent to be exposed to English at least 70% of the time, and to have no known

language, communication, or uncorrected hearing or vision problems. Data from 6 additional children were excluded, 2 due to failure to point correctly on at least one of two pointing training trials, or failure to point at all on at least 1 of the 3 experimental trials; and 4 due to excessive track loss (defined as more than 65% track loss during Familiarization or Test of all trials).

Stimuli

We used a new set of stimuli to see if the results from Experiment 1 would generalize; they were modeled after Experiment 1, with a few minor differences mentioned below, and consisted of two pointing training trials and three experimental trials. The visual stimuli for training trials were video clips of familiar animals (e.g., cat), and auditory prompts were prerecorded sentences like "where's the cat?" For experimental trials, as in Experiment 1, the videos were recordings of live actors (e.g., a boy waving) or objects manipulated by a hand (e.g., a ball moving up and down). Auditory stimuli were recorded by a female native English speaker using child-directed speech and edited in Praat (Boersma & Weeknink, 2014) and combined with the visual stimuli in Final Cut Pro X.

Apparatus and procedure

The procedure was identical to Experiment 1, except that we did not collect vocabulary data because no MCDI was appropriate for the full age range, and because Experiment 1 did not yield significant effects of vocabulary. Children also participated in a different verb learning task (with transitive verbs, and no manipulation of modification) either before or after their participation in this experiment.

Pointing training
In the pointing training session, children were given two trials; in each, children saw two familiar animals side-by-side and were asked to point to one (e.g., "Where's the cat?").

Experimental trials
As in Experiment 1, children were randomly assigned to one of three conditions in a between-subject design (each N = 20): Light, Heavy-Unnecessary, or Heavy-Necessary. The trials were structured similarly to Experiment 1, and they differed by condition only in the Familiarization phase. See Table 6 for an example trial and Table 7 for a summary of stimuli in all trials.

Table 6. An example trial (Experiment 2).

Table 7. A summary of stimuli used in all trials in Experiment 2.

Trial	Actors		Actions	Sentence
	Light & Heavy-Unnecessary conditions	Heavy-Necessary condition		
1	round ball, square box	round ball, football	move up/down, rotate	"The (round) ball is *kradd*ing"
2	tall boy, short girl	tall boy, short boy	wave, clap	"The (tall) boy is *pilk*ing"
3	standing dog, sitting pig	standing dog, sitting dog	kick foot, raise arms	"The (standing) dog is *gorp*ing"

Familiarization phase. The trial structure differed minimally from Experiment 1. We made one principal change: Because we wanted to assess children's eye gaze as they were in the process of assigning meaning to the novel verb, we slightly changed the Familiarization phase, now presenting the linguistic stimulus during a blank screen, without distraction from the visual scenes. We made this change because children's gaze to dynamic scenes is not straightforwardly in response to processing of co-occurring language, but may also be in response to aspects of motion in the scenes (Valleau, Konishi, Golinkoff, Hirsh-Pasek, & Arunachalam, 2018). Therefore, this design choice allowed us to assess children's eye gaze when the scenes reappeared, as they sought to assign meaning to the novel verb. Further, because the Light condition differed from the Heavy conditions in the length of the subject DP, this design choice also ensured that we were assessing gaze in both conditions after they had heard the entire auditory stimulus. Although we measured eye gaze after the sentence was complete, the sentence and the assessment window were very close in time. The offset of the sentence was aligned to the onset of the assessment window, and the mean duration between the onset of the novel verb and the onset of the assessment window was 0.65 s in the Light condition and 0.61 s in the Heavy conditions. Given that Fernald, Thorpe, and Marchman (2010) found that 30- and 36-month-olds required approximately 0.5 s to look to a target named with an adjective+noun phrase, we think that children's gaze responses during familiarization were relatively closely tied to their process of interpreting the sentence – that is, related to their abilities to parse the sentence incrementally. There were other minor differences in the timeline of the trial from Experiment 1, principally due to differences in the duration of auditory stimuli.

Test phase. As in Experiment 1, except for minor differences in the trial timeline as shown in Table 6.

Coding

Coding of children's pointing behavior at test was undertaken as for Experiment 1, except that reliability coding by a second experimenter occurred immediately after the session by a second experimenter who watched the recording.[9] Disagreements were resolved by discussion immediately after the session. As in Experiment 1, the first point (if more than one) produced on each trial served as the dependent measure: 0 for pointing to the distractor, and 1 for the target. The 3% (5 out of 180) of trials on which children did not point were excluded from analysis. Note that we did not evaluate gaze data at test because we wanted the results to be directly comparable to the results from Experiment 1, and children's pointing behavior can interfere with evaluating their gaze if they block the eye-tracker with their hand when executing a point.

For gaze data during familiarization, 3% (6 out of 180) of trials that had excessive track loss (i.e., more than 65% of that trial's Familiarization or Test phase were track loss) were excluded. In analyzing the data, we excluded frames that had track loss, and coded frames on which children looked to the target as 1, and all other frames as 0 (including frames on which they looked to the

---

[9] This was to reduce confusion caused by the fact that in our setup, the experimenter behind the curtain had a mirror-image perspective of the child's behavior.

distractor, as well as frames for which gaze coordinates were captured by the eye-tracker but were outside any area of interest).

Results

Pointing data at test. As in Experiment 1, we analyzed children's pointing at test to see if they learned the verb and if performance was related to the type of subject DP. The results are illustrated in Figure 2. Mean pointing score in the Light condition was 0.67 ($SD = 0.26$), in the Heavy-Unnecessary condition, 0.57 ($SD = 0.22$), and in the Heavy-Necessary condition, 0.48 ($SD = 0.20$). In the Light condition, 13 children chose the target on at least two trials; 6 of these did so on all three trials. In the Heavy-Unnecessary condition, 14 chose the target on at least two trials, but only 1 did so on all three trials. And in the Heavy-Necessary condition, 9 succeeded on at least two trials, but none did so on all three. As in Experiment 1, we adopted a significance level of 0.05 for all statistical analyses. Analyses were carried out in R (version 3.3.3) using the lme4 package (version 1.1–12).

As in Experiment 1, we first submitted pointing data from each condition to a mixed-effects logistic regression model, which allowed us to add participant and trial as random factors and age (in months; centered around its mean) as a continuous predictor. We compared pointing responses to chance by evaluating the intercept parameter. The results replicated those in Experiment 1: Only the Light condition yielded above-chance learning (intercept parameter = 0.81, $z = 2.02$, $p < .05$); the Heavy-Unnecessary condition (intercept parameter = 0.34, $z = 1.14$, $p = .26$) and the Heavy-Necessary condition (intercept parameter = -0.071, $z = -0.22$, $p = .82$) did not. There was no effect of age in any condition, despite the wider age range. We also entered data in all conditions to a single mixed-effects logistic regression model, with condition as fixed effect, participant and trial as random effects, and age (in month; centered around its mean) as a continuous predictor. Unlike in Experiment 1, here, the difference between the Light and the Heavy-Necessary condition was statistically significant ($z = 2.24$, $p < .05$); specifically, the Light condition's mean performance score was 0.19 higher than that in the Heavy-Necessary condition (Cohen's $d = 0.84$). The differences between the Light and the Heavy-Unnecessary conditions ($z = 0.87$, $p = .39$) and between the
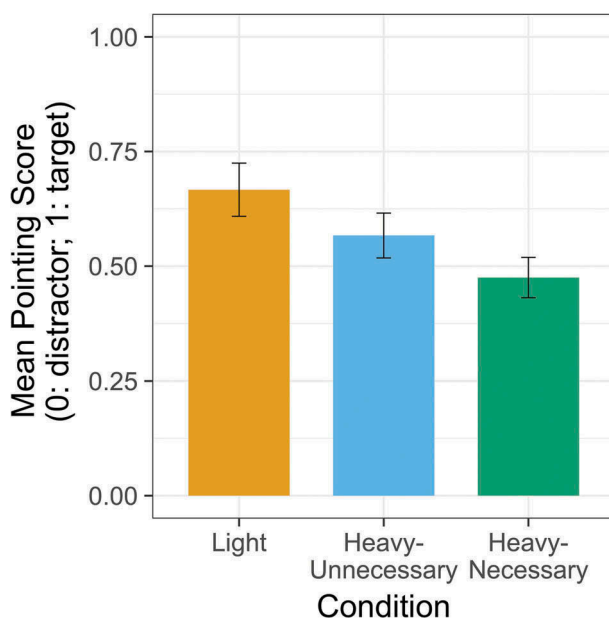
Figure 2. Mean pointing score in Experiment 2 (Error bars indicate standard errors of participant means).

Table 8. Summary of by-trial pointing performance in Experiment 2.

	Light		Heavy-Unnecessary		Heavy-Necessary	
	Mean	SD	Mean	SD	Mean	SD
Trial 1	0.67	0.49	0.74	0.45	0.53	0.51
Trial 2	0.85	0.37	0.53	0.52	0.60	0.50
Trial 3	0.50	0.51	0.45	0.51	0.30	0.47

two Heavy conditions ($z = 1.33$, $p = .18$) were not significant. There was no effect of age. See Table 8 for a summary of performance by trial.

Gaze data during familiarization

To investigate the relation between children's processing of subject DPs and verb learning, we analyzed children's eye gaze during Familiarization as an index of their learning process. The Familiarization phase provided children 6 seconds to look at the two scenes after hearing the critical auditory stimulus (e.g., "The (tall) boy is *pilk*ing"). We examined the first 2.5 seconds of this 6-second period because this is the longest time window used both in studies of modified DP processing (e.g., Fernald et al., 2010) and studies of novel verb learning (e.g., Arunachalam, Escovar, Hansen, & Waxman, 2013). Although the children in our study were older on average and should be faster to direct their attention to the relevant scene, we used a 2.5-second window rather than artificially restricting its duration.

First, to assess whether children successfully identified the target event during Familiarization, we compared their mean target looks over the 2.5-second window to chance using a one-sample *t*-test. To ensure that chance was truly 0.5, we only included data points that had either a target look or a distractor look (in other words, looks to other areas of the screen were removed from this analysis). Results indicated that children in all conditions attended to the target scene significantly more than chance: on average, children in the Light condition looked at the scene depicting the target event 81.3% of the time ($SD = 13.4$%; $t = 10.42$, $p < .001$), in the Heavy-Unnecessary condition, 69.0% of the time ($SD = 16.9$%; $t = 5.01$, $p < .001$), and in the Heavy-Necessary condition, 62.0% of the time ($SD = 22.6$%; $t = 2.38$, $p < .05$). These results are illustrated in Figure 3.

Although children preferred the target scene to the distractor scene in all three conditions, this does not mean that they did so to the same degree, or along the same time course, across conditions. We next used Growth Curve Analysis (Mirman, 2014) to analyze the time course of target fixation. This analysis included looks that were directed to neither the target nor distractor, coded as 0. Since the dependent variable was categorical (i.e., fixating the target or not), we applied the empirical logit transformation (Barr, 2018), which is robust to values at or near the boundaries (0 and 1). The overall time course of target looking was captured with a second-order orthogonal polynomial, with a fixed effect of condition and random effect of participant on both time terms, and a random effect of participant-by-condition on the linear time term.[10] The model was run twice to obtain all pair-wise comparisons among the three conditions, first with the Light condition as the reference level, and second with the Heavy-Necessary condition as the reference level. Statistical significance for individual parameter estimates was assessed using the normal approximation (i.e., treating the *t*-value as a *z*-value).

The data and model fits are shown in Figure 4. There was a significant difference between the Light condition and the Heavy-Unnecessary condition ($t = -2.18$, $p < .05$), and also between the Light condition and the Heavy-Necessary condition ($t = -2.99$, $p < .01$), but the difference between the two Heavy conditions was not significant ($t = 0.81$, $p = .42$). These results indicated overall higher fixation probability for the Light condition than either of the two Heavy conditions. There was also an effect of the Heavy-Necessary condition on the quadratic time term, both in comparison to the Light condition ($t = -3.33$, $p = .001$) and to the Heavy-Unnecessary condition ($t = -3.16$,

---

[10]The model failed to converge with a participant-by-condition random effect on both time terms, so this random effect on the quadratic time term was removed.
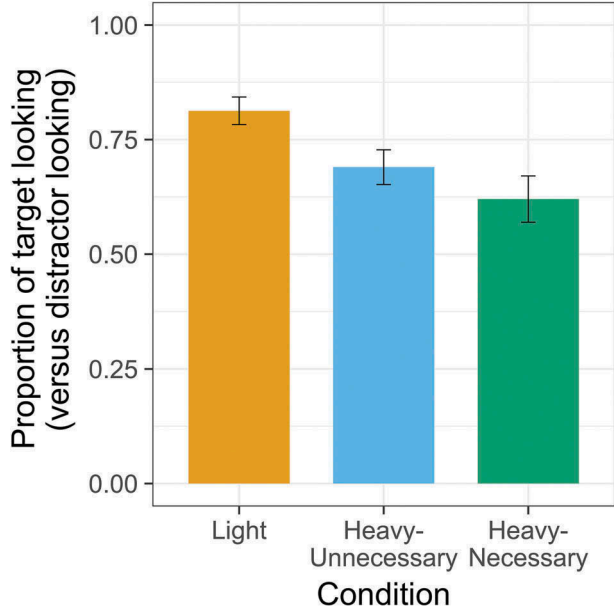
Figure 3. Mean proportion of target looks during Familiarization Phase in Experiment 2 (Error bars indicate standard errors of participant means).
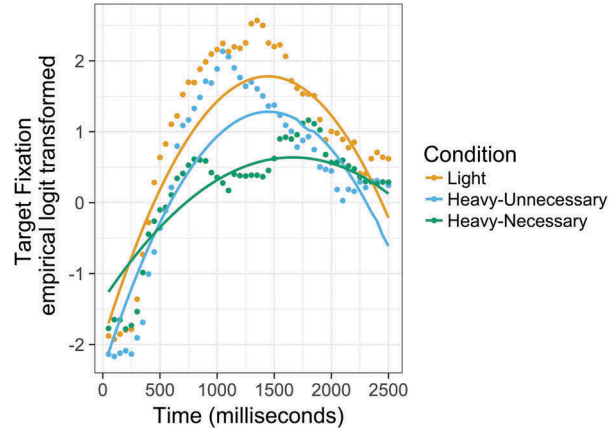
Figure 4. Fixation time course during Familiarization Phase in Experiment 2 and growth curve model fit.

$p < .01$), suggesting that the rate of change in the Heavy-Necessary condition was different from that in the other two conditions. No other significant effects obtained.

Discussion

These results addressed the two goals of Experiment 2. First, the trend in Experiment 1—that modified subject DPs might have hindered, rather than facilitated, verb learning—was replicated in Experiment 2 with a different set of stimuli, and with children in a broader age range. Importantly, the cross-condition comparison showed a statistically significant difference in

Experiment 2 between the Light and the Heavy-Necessary conditions. Second, children in all conditions preferred the target event to the distractor event during Familiarization, suggesting that they were able to process the subject DPs, even if they included a modifier.

However, even though children in the two Heavy conditions were able to find the referent of the modified DP (indicated by above-chance target fixation), their overall probability of target fixation was smaller in comparison to that in the Light condition (indicated by between-condition differences from the Growth Curve Analysis). This suggests that the impact of modified DPs on children's verb learning may have begun in the Familiarization phase; however, the consequence was not failure to find the DP's referent, but perhaps only the efficiency with which the referent was identified. Second, we found no evidence for a pragmatic account, that an unnecessary modifier (as in the Heavy-Unnecessary condition) would have caused more difficulty than an appropriately informative modifier (as in the Heavy-Necessary condition). (If anything, although there were no statistically significant differences between the two Heavy conditions, children looked less to the target in the Heavy-Necessary condition.) This new gaze data from the Familiarization phase supports our interpretation of the data in Experiment 1, that at least in our task, children's difficulty with verb learning in contexts with modified subject DPs is related to processing load but not to pragmatic felicity.

General discussion

Although prior work has established that a novel verb's linguistic context provides useful information that supports acquisition of its meaning, less is known about what makes a linguistic context more or less supportive. In particular, for young children seeking to acquire the meaning of a new verb, is more information in the linguistic context necessarily better than less? In two experiments, we provide evidence that more information can interfere with young children's verb learning outcomes (as seen in Experiment 1 and 2). Further, this interference is evident even during the verb learning process (as seen in Experiment 2).

In Experiment 1, preschoolers showed above-chance performance with novel intransitive verbs when the subject was an unmodified lexical DP (e.g., "the girl"), but were at chance when the subject was modified (e.g., "the tall girl"). Experiment 2 replicated this finding, also demonstrating a significant difference between conditions, and further showed that children's failure with a modified DP was *not* because they could not parse the DP and identify its referent. What, then, underlies the interference of a modified DP? Our findings are inconsistent with a simple pragmatics account under which children struggle with modifiers that are unnecessary, because children also struggled when the modifier was necessary to identify the referent. Further, there were no significant differences between the Heavy-Unnecessary condition (in which the modifier was overinformative) and the Heavy-Necessary condition (in which it was appropriately informative). The lack of a pragmatic effect may not be surprising because children up to 5 years of age are tolerant of overinformative statements like "find the bird with the feather" when there is only one bird present (Davies & Katsos, 2010; Morisseau et al., 2013).[11]

Instead of pragmatic felicity, our results suggest that the explanation is likely due to children's limited processing abilities. Prior work has shown how limitations on the developing parser influence language comprehension, in children's abilities to revise an initial interpretation of garden-path sentences (Choi & Trueswell, 2010) and passive sentences (Huang & Arnold, 2016). Our experiments did not require revision, but instead posed a heavy processing load that apparently hindered verb learning.

---

[11]Although there were no statistically significant differences between the Heavy-Necessary condition and the Heavy-Unnecessary condition, trends suggested that the former may have been more difficult. If further work suggests that this is true, this could be because children could not simply "listen through" the modifier (Thorpe & Fernald, 2006) but instead were required to parse and integrate it, and/or because the visual scenes in the Heavy-Necessary condition (e.g., two boys) were more similar to each other than in the Heavy-Unnecessary condition (e.g., boy and girl), potentially making it more difficult to distinguish them (Arias-Trejo & Plunkett, 2010; Bergelson & Aslin, 2017).

Interestingly, it was not the case that children were unable to process the modified DPs at all. Results from the familiarization phase in Experiment 2 clearly indicated that children were able to find the referents of modified DPs, but nonetheless still failed to learn the novel verbs downstream. These findings shed light on the mechanism by which modified DPs hinder learning. We suggest that processing the modified DPs exhausted children's cognitive resources for language processing – their abilities to engage in processes such as lexical access, structure building, and information storage – thereby leaving inadequate resources for positing a representation for the novel verb. That is, children successfully processed the modified DPs, but doing so taxed their parsers enough that they had insufficient resources left to fulfill the other steps required for learning the verb.

Similarly, in a novel noun learning task, Fernald et al. (2008) found that 3-year-olds who more quickly processed modified DPs were more likely to learn the novel nouns that came later (e.g. "There is a blue cup on the *deebo*"). In their study, noun learning outcomes were a function of individual children's processing speed; in our study, verb learning outcomes were a function of the processability of the linguistic context. Children in that study were more successful overall than in our study, but this is perhaps not surprising given that verb learning is a more complex task than noun learning, both in general (e.g., Gleitman et al., 2005) and in our task.

Processing limitations are also implicated in children's spontaneous production as well as in adults' language processing. In children's spontaneous production, it has been observed that they are more likely to reduce the subject to a shorter form (e.g. reducing a full lexical DP to a pronoun) or completely omit it in sentences with longer VPs (than in sentences with shorter VPs) (e.g., Bloom, 1970; 1990), and in sentences with certain elements such as negations and particles (than without) (e. g., Bloom, 1970; Bloom, Miller, & Hood, 1975). Similar phenomena have been observed in isolated deaf children's sign production (Feldman, Goldin-Meadow, & Gleitman, 1978). On a processing account, children's limited processing ability leads to a tendency to produce shorter sentences, and reducing or omitting subjects may be one way children that cope with this processing bottleneck (e. g., Bloom, 1970; 1990). Quite conceivably, the same processing bottleneck that affects language production may also influence language learning.

In adult language processing, it has also been robustly documented that referential expressions with more information slow down processing if the extra information is already active in the discourse, and processing is fastest when there is a balance between the referential expression's discourse function and processing cost (e.g., Almor, 1999; Almor, Arunachalam, & Strickland, 2007; Alonso-Ovalle, Fernández-Solera, Frazier, & Clifton, 2002; Ariel, 1991; Filipović & Hawkins, 2013; Givón, 1983; Gordon, Grosz, & Gilliom, 1993; Gundel, Hedberg, & Zacharski, 1993). For example, sentences with anaphors carrying information that is repeated from their antecedents are read more slowly than sentences with pronominal anaphors (e.g., Gordon et al., 1993). Speakers' choices of referential expressions in production also are influenced by both informativity and processability— usually, a low-cost referring expression (e.g., a pronoun) is preferred; but when the referent is harder to identify, more information is needed, which may warrant the use of a high-cost referring expression (e.g., a lexical DP) (see Almor & Nair, 2007 for review). Our results are therefore in line with accounts of language production and comprehension more generally.

Several questions remain. First, if more information incurs a greater processing load, what aspect of the linguistic information incurs processing load—phonological form, semantic content, syntactic structure, or some combination of these? The current experiments were not designed to pit one factor against another—the modified DPs were "heavier" than the unmodified DPs both semantically and phonologically, and were also syntactically more complex. Future work may tease these apart by, for example, contrasting DPs that are similar in semantic and syntactic complexity but vary in phonological complexity (e.g., "butterfly" vs. "bee").

Second, we have argued that learners' processing ability limits their use of linguistic context, but we have not defined what constitutes "processing ability", though we suspect it is some combination of factors like verbal working memory, speed of lexical access, and speed of structure building. Experience with processing a phrase may alleviate the burdens associating with processing a complex

linguistic stimulus. For example, younger children in Lidz et al.'s (2009) experiment failed to learn a verb preceded by a lexical DP (e.g., "The flower is *blick*ing"), but succeeded if they were exposed to the DP beforehand (e.g., "Look at the flower!", "Here's a flower!"). Similarly, in the current study, there is a suggestive trend, though it is not statistically reliable, that children may have improved across trials, also consistent with the notion that practice eases processing.

Although we believe the results of the current experiments are best interpreted with a processing account, we do not mean to suggest that processing is the only relevant aspect of the verb learning challenge. While more information is not always better, children do need a certain minimal amount of semantic information in verb learning tasks (e.g., Arunachalam & Waxman, 2011, 2015; Imai et al., 2008). Taken together, we think the prior verb learning results, combined with those of the current study, highlight the fact that informativity and processability of the linguistic context must achieve a balance. How exactly this balance is achieved will depend on many factors including the type of event being labeled and what other candidate interpretations are available (see also He & Arunachalam, 2017 for review).

In future work, we will also explore alternate ways of providing informative contexts that do not tax the parser so much as to hinder word learning – after all, in some situations modifiers may be necessary to disambiguate referents. It may be, for example, that information encountered *after* a novel verb can be more easily integrated than information in subject position. In support of this hypothesis, Syrett et al. (2014) found that semantically informative post-verbal adverbs can aid verb learning, and Arunachalam (2016) found that DPs with postnominal adjectival modifiers are more easily processed than those with prenominal modifiers. Of course, processing load can be contributed by elements other than prenominal modifiers – we are more broadly suggesting that different ways of "packaging" informative elements may result in the same amount of information being conveyed with a lightened processing load.

Three methodological features of our experiments are also worth noting. First, we introduced the novel verb in a complex visual context with two simultaneous events, only one of which is labeled by the verb to more closely resemble the cluttered environment in real-world learning. Although several verb learning studies investigating syntactic bootstrapping have used similarly complex visual contexts, for which linguistic context was required to identify the intended referent, in those studies the visual scenes were disambiguated by argument structure (specifically transitive vs. intransitive syntax) (e.g., Naigles, 1990; Naigles & Kako, 1993). In the current study, the verb's argument structure was held constant, and we instead manipulated the amount of semantic content available in the verb's subject argument. Second, unlike prior verb learning studies which only examined verb learning outcomes by evaluating children's performance in a "test," we also evaluated children's eye gaze while they were in the process of assigning a meaning to the novel verb (see also Valleau & Arunachalam, 2017). Third, in our design children were only exposed to the novel verb once, in one linguistic context, before being tested. That they succeeded in even one such experimental condition is a strong testament to their abilities to acquire new verbs, provided they are given a supportive linguistic context.

Our results inform a broad question that is of theoretical and practical importance: What constitute optimal linguistic contexts for verb learning? In particular, while informativity is undoubtedly beneficial, for young children the most supportive linguistic context must also be processable. This laboratory study has implications for thinking about how children may learn new verbs from the input they encounter in daily life, and may have implications for treatment for children with Developmental Language Disorder, who struggle with verb acquisition and also have difficulty comprehending modified DPs (Leonard, Deevy, Fey, & Bredin-Oja, 2013). As Givón (1985) puts it, a certain kind of input may be very frequent, but "is nevertheless *irrelevant* to what the child is *likely* to be able to absorb at that particular point in his/her cognitive and linguistic development" (p. 1007). Studies of the input, then, must be integrated with studies of children's abilities to process and learn from linguistic context in order to provide a full picture of their intake of that input.

Acknowledgments

Disclosure statement

Funding

References

Almor, A. (1999). Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review*, *106*(4), 748–765. doi:10.1037/0033-295X.106.4.748

Almor, A., Arunachalam, S., & Strickland, B. (2007). When the creampuff beat the boxer: Working memory, cost, and function in reading metaphoric reference. *Metaphor and Symbol*, *22*(2), 169–193. doi:10.1080/10926480701235478

Almor, A., & Nair, V. A. (2007). The form of referential expressions in discourse. *Language and Linguistics Compass*, *1* (1–2), 84–99. doi:10.1111/lnco.2007.1.issue-1-2

Alonso-Ovalle, L., Fernández-Solera, S., Frazier, L., & Clifton, C. (2002). Null vs. overt pronouns and the topic-focus articulation in Spanish. *Italian Journal of Linguistics*, *14*(2), 151–169.

Arias-Trejo, N., & Plunkett, K. (2010). The effects of perceptual similarity and category membership on early word-referent identification. *Journal of Experimental Child Psychology*, *105*(1–2), 63–80. doi:10.1016/j.jecp.2009.10.002

Ariel, M. (1991). The function of accessibility in a theory of grammar. *Journal of Pragmatics*, *16*(5), 443–463. doi:10.1016/0378-2166(91)90136-L

Arunachalam, S. (2016). A new experimental paradigm to study children's processing of their parent's unscripted language input. *Journal of Memory and Language*, *88*, 104–116. doi:10.1016/j.jml.2016.02.001

Arunachalam, S., Escovar, E., Hansen, M. A., & Waxman, S. R. (2013). Out of sight, but not out of mind: 21-month-olds use syntactic information to learn verbs even in the absence of a corresponding event. *Language and Cognitive Processes*, *28*(4), 417–425. doi:10.1080/01690965.2011.641744

Arunachalam, S., Leddon, E. M., Song, H., Lee, Y., & Waxman, S. R. (2013). Doing more with less: Verb learning in Korean-acquiring 24-month-olds. *Language Acquisition*, *20*(4), 292–304. doi:10.1080/10489223.2013.828059

Arunachalam, S., & Waxman, S. R. (2011). Grammatical form and semantic context in verb learning. *Language Learning and Development*, *7*(3), 169–184. doi: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. doi:10.1080/10489223.2014.928300

Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, *47*(1), 31–56. doi:10.1177/00238309040470010201

Backscheider, A. G., & Shatz, M. (1993). Children's acquisition of the lexical domain of color. *The 29th Annual Proceedings of the Chicago Linguisitc Society*. Chicago.

Barr, D. (2018). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, *59*(4), 457–474. doi:10.1016/j.jml.2007.09.002

Behrend, D. A. (1990). The development of verb concepts: Children's use of verbs to label familiar and novel events. *Child Development*, *61*(3), 681–696.

Bergelson, E., & Aslin, R. N. (2017). Nature and origins of the lexicon in 6-mo-olds. *Proceedings of the National Academy of Sciences*, *114*(49), 12916–12921. doi:10.1073/pnas.1712966114

Bloom, L. (1970). Language development: Form and function in emerging grammars. In *MIT research monograph series, No 59*. The MIT Press. doi:10.7916/D8S75GT1

Bloom, L., Miller, P., & Hood, L. (1975). Variation and reduction as aspects of competence in language development. In A. Pick (Ed.), *The 1974 minnesota symposium on child psychology* (Vol. 9, pp. 3–55). Minneapolis: University of Minnesota Press.

Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry*, *21*, 491–504.

Boersma, P., & Weeknink, D. (2014). *Praat version 5.3: Doing phonetics by computer*. Retrieved from http://www.fon.hum.uva.nl/praat/

Choi, Y., & Trueswell, J. C. (2010). Children's (in)ability to recover from garden-paths in a verb-final language: Evidence for developing control in sentence. *Journal of Experimental Child Psychology*, *106*(1), 41–61. doi:10.1016/j.jecp.2010.01.003

Chomsky, N. (1959). A review of BF Skinner's verbal behavior. *Language*, *35*(1), 26–58. doi:10.2307/411334

Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*(pp. 1–33). Hillsdale, NJ: Lawrence Erlbaum Associates.

Davies, C., & Katsos, N. (2010). Over-informative children: Production/comprehension asymmetry or tolerance to pragmatic violations? *Lingua*, *120*(8), 1956–1972. doi:10.1016/j.lingua.2010.02.005

Feldman, H., Goldin-Meadow, S., & Gleitman, L. R. (1978). Beyond Herodotus: The creation of language by isolated deaf children. In J. Locke (Ed.), *Action, gesture and symbol: The emergence of language*(pp. 351–414). New York, NY: Academic Press.

Fernald, A., Marchman, V. A., & Hurtado, N. (2008). Input affects uptake: How early language experience influences processing efficiency and vocabulary learning. *7th IEEE International Conference on Development and Learning* (pp. 37–42). Monterey, CA. doi:10.1109/DEVLRN.2008.4640802

Fernald, A., Thorpe, K., & Marchman, V. (2010). Blue car, red car: Developing efficiency in online interpretation of adjective-noun phrases. *Cognitive Psychology*, *60*(3), 190–217. doi:10.1016/j.cogpsych.2009.12.002

Filipović, L., & Hawkins, J. A. (2013). Multiple factors in second language acquisition: The CASP model. *Linguistics*, *51*(1), 145–176. doi:10.1515/ling-2013-0005

Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive Psychology*, *31*(1), 41–81. doi:10.1006/cogp.1996.0012

Fisher, C. (2002). The role of abstract syntactic knowledge in language acquisition: A reply to Tomasello (2000). *Cognition*, *82*(3), 259–278. doi:10.1016/S0010-0277(01)00159-7

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

Forbes, J. N., & Farrar, M. J. (1993). Children's initial assumptions about the meaning of novel motion verbs: Biased and conservative? *Cognitive Development*, *8*(3), 273–290. doi:10.1016/S0885-2014(93)80002-B

Forbes, J. N., & Farrar, M. J. (1995). Learning to represent word meaning: What initial training events reveal about children's developing action verb concepts. *Cognitive Development*, *10*(1), 1–20. doi:10.1016/0885-2014(95)90016-0

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

Givón, T. (1983). *Topic continuity in discourse: A quantitative cross-language study* (Vol. 3). Philadelphia, PA: John Benjamins Publishing.

Givón, T. (1985). Function, structure and language acquisition. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 2, pp. 1005–1027). Hillsdale, N.J.: Erlbaum Associates.

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

Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development*, *1*(1), 23–64. doi:10.1207/s15473341lld0101_4

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. doi:10.1207/s15516709cog1703_1

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics 3: Speech acts* (pp. 41–58). New York, NY: Academic Press.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, *69*(2), 274–307. doi:10.2307/416535

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

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.

Hesketh, L. J., & Ellis Weismer, S. (1997). Do preschoolers evidence mutual exclusivity bias in verb-synonym learning. *Poster Presented at the American Speech-Language Hearing Association Annual Convention*. Boston, MA.

Hirsh-Pasek, K., & Golinkoff, R. M. (1996). *The origins of grammar: Evidence from early language comprehension*. Cambridge, MA: MIT Press.

Huang, Y. T., & Arnold, A. R. (2016). Word learning in linguistic context: Processing and memory effects. *Cognition*, *156*, 71–87. doi:10.1016/j.cognition.2016.07.012

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. doi: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. doi:10.1111/j.1467-8624.2008.01171.x

Jaeger, T. F., & Levy, R. P. (2007). Speakers optimize information density through syntactic reduction. In B. Schölkopf, J. C. Platt, & T. Hoffman (Eds.), *Advances in Neural Information Processing Systems 19 (NIPS 2006)* (pp. 849–856). Cambridge, MA: The MIT Press. doi:10.7551/mitpress/7503.003.0111

Kersten, A. W., & Smith, L. B. (2002). Attention to novel objects during verb learning. *Child Development*, *73*(1), 93–109. doi:10.1111/1467-8624.00394

Kidd, E., Bavin, E. L., & Brandt, S. (2013). The role of the lexicon in the development of the language processor. In D. Bittner & N. Ruhlig (Eds.), *Lexical bootstrapping: The role of lexis and semantics in child language*(pp. 217–244). Berlin, Germany: De Gruyter Mouton.

Landau, B., & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.

Leonard, L. B., Deevy, P., Fey, M. E., & Bredin-Oja, S. L. (2013). Sentence comprehension in specific language impairment: A task designed to distinguish between cognitive capactiy and syntactic complexity. *Journal of Speech Language and Hearing Research*, *56*(2), 577–589. doi:10.1044/1092-4388(2012/11-0254)

Lidz, J., Bunger, A., Leddon, E., Baier, R., & Waxman, S. R. (2009). *When one cue is better than two: Lexical vs. syntactic cues to verb meaning*. Unpublished manuscript.

Markman, E. M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 72–106). New York, NY: Cambridge University Press.

Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, *16*(1), 1–27. doi:10.1016/0010-0285(84)90002-1

Merriman, W., Marazita, J., & Jarvis, L. (1995). Children's disposition to map new words onto new referents. In M. Tomasello & W. Merriman (Eds.), *Beyond names for things: Young children's acquisition of verbs* (pp. 147–183). New York, NY: Psychology Press.

Mirman, D. (2014). *Growth curve analysis and visualization using R* (1st ed.). New York, NY: Chapman and Hall/CRC.

Morisseau, T., Davies, C., & Matthews, D. (2013). How do 3- and 5-year-olds respond to under- and over-informative utterances? *Journal of Pragmatics*, *59*, 26–39. doi:10.1016/j.pragma.2013.03.007

Naigles, L. R. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*(2), 357–374. doi:10.1017/S0305000900013817

Naigles, L. R., & Kako, E. T. (1993). First contact in verb acquisition: Defining a role for syntax. *Child Development*, *64* (6), 1665–1687.

Newport, E. L. (1975). *Motherese: The speech of mothers to young children* (Doctoral Dissertation). University of Pennsylvania. Retrieved from https://repository.upenn.edu/dissertations/AAI7524108

Omaki, A., & Lidz, J. (2015). Linking parser development to acquisition of syntactic knowledge. *Language Acquisition*, *22*(2), 158–192. doi:10.1080/10489223.2014.943903

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

Snow, C. E. (1972). Mothers' speech to children learning language. *Child Development*, *43*, 549–565. doi:10.2307/1127555

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. doi:10.1080/15475441.2013.840493

Thorpe, K., & Fernald, A. (2006). Knowing what a novel word is not: Two-year-olds 'listen through' ambiguous adjectives in fluent speech. *Cognition*, *100*, 389–433. doi:10.1016/j.cognition.2005.04.009

Trueswell, J. C., & Gleitman, L. R. (2007). Learning to parse and its implications for language acquisition. In G. Gaskell & G. T. M. Altmann (Eds.), *Oxford handbook of psycholinguistics* (pp. 635–657). Oxford, England: Oxford University Press.

Trueswell, J. C., Sekerina, I., Hill, N. M., & Logrip, M. L. (1999). The kindergarten-path effect: Studying on-line sentence processing in young children. *Cognition*, *73*, 89–134.

Valleau, M., & Arunachalam, S. (2017). The effects of linguistic context on visual attention while learning novel verbs. In M. LaMendol & J. Scott (Eds.), *Proceedings of the 41st annual Boston University Conference on Language Development (BUCLD)* (pp. 691–705). Somerville, MA: Cascadilla Press.

Valleau, M., 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. doi:10.1002/adsc.201

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

Wagner, K., Dobkins, K., & Barner, D. (2012). Color word learning is a gradual inductive process. In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *34th annual meeting of the cognitive science society* (pp. 1096–1101). Austin, TX: Cognitive Science Society.

Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: Simple structural cues to verb meaning. *Child Development*, *83*(4), 1382–1399. doi:10.1111/j.1467-8624.2012.01783.x