Quantifying the syntactic bootstrapping effect in verb learning through meta-analysis

Anjie Cao[1] & Molly Y. Lewis[2]

[1] Department of Psychology, Stanford University

[2] Department of Psychology, Carnegie Mellon Unviversity

Author Note

okie

Correspondence concerning this article should be addressed to Anjie Cao, what is this. E-mail: acao@andrew.cmu.edu

Abstract

so abstract!

Keywords: language acquisition, syntactic bootstrapping, meta-analysis

Word count:

Quantifying the syntactic bootstrapping effect in verb learning through meta-analysis

Introduction

In order to become fluent users of language, children must learn the names not only for concrete nouns, like “ball” and “apple”, but also verbs like “eat” and “throw”. Learning the meaning of a verb presents a particular challenge to learners because the perceptual information about a verb’s meaning is complex – it is ephemeral and temporally dynamic. How are children able to overcome this perceptual ambiguity in order to learn verb meanings? One of the central theoretical proposals is that children use syntactic information in a verb’s linguistic context to infer its meaning (Gentner, 2006; Gentner & Boroditsky, 2001), a phenomenon known as “syntactic bootstrapping” (Brown, 1957; Gleitman, 1990; Landau & Gleitman, 1985; Naigles, 1990). For example, if a child hears the sentence “Mom pilked the apple”, the child can use knowledge about the language’s syntax to infer that “pilked” refers to an event where “Mom” is an agent acting on a patient, “the apple”, thereby constraining the hypothesis space of possible referents. A large body of empirical literature over the past thirty years has experimentally tested children’s ability to use syntactic information to infer verb meanings. The goal of the current paper is to synthesize this literature quantitatively using meta-analytic methods in order to evaluate the evidential value of this literature and understand theoretical and methodological moderators of the strength of the effect.

The earliest work on syntactic bootstrapping demonstrated that children are able to use coarse-grain syntactic information to constrain hypotheses about a word’s meaning. Broad syntactic categories, like noun and verb, are probabilistically linked to semantic categories, like concrete things and actions. Brown (1957) demonstrated that children are able to make use of this information by presenting children with sentences containing a novel word in different syntactic categories (e.g., “Do you know what it means to sib?” (verb) vs. “Do you know what a sib is?” (noun)), and measuring whether children inferred the novel word to refer to an action or a concrete entity. Subsequent studies demonstrated that children are not only able to use broad syntactic categories to infer word meaning, they are also able to make use of more nuanced syntactic information. In particular, the syntax of a sentence provides information about the number of arguments a predicate takes. For example, verbs denoting a causative event (e.g. pushing) tend to take two noun arguments (a "pusher" and a "pushee"), whereas verbs denoting self-generated motion (e.g., waving) take only one noun argument (a "waver"). A large body of work has demonstrated that children are able to successfully map a predicate with two noun arguments (a transitive predicate) to a two-agent causative event, and a predicate with one noun argument (an intransitive predicate) to a non-causative event.

The ability to infer a novel verb meaning on the basis of the number of preicates in a sentence is typically tested in a paradigm where children are simultaneously presented with two visual stimuli and an auditory sentence (Fernald, Pinto, Swingley, Weinbergy, & McRoberts, 1998; Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987). Canonically (e.g., Naigles, 1990), one visual stimulus depicts two characters doing an unfamiliar causative action to each other (e.g., a novel pushing motion), and the other depicts one or more characters doing an unfamiliar non-causative action (e.g., a novel waving gesture). Children then hear a transitive (e.g., "The duck is gorping the bunny") or intransitive sentence (e.g., "The duck and the bunny are gorping") with a novel verb, and then asked to find the novel verb ("Where's the gorping now?"). Children's fixation time or pointing selections are then measured. Evidence for syntactic bootstrapping is found when children look longer at the visual stimulus matching the syntactic structure of the sentence.

The body of empirical literature investigating syntactic bootstrapping has demonstrated that children are able to successfully use syntactic information to infer verb meaning given a range of linguistic and visual stimuli, and with many different samples of children in a range of experimental manipulations. However, while these studies provide a binary description of the effect ("children can bootstrap meanings from sytnax"), they do not quantify the *strength* of the effect. Quantifying the strength of the effect allows researchers to assess the degree to which a learning strategy can be used by young learners to acquire verb meanings. Further, it allows for the quantitative comparison to other proposed learning mechanisms for inferring the meaning of a word in a local context. For example, children may be able to infer a word's meaning by using co-occurrence statistics between words and referents ("cross-situational learning"; Smith & Yu, 2008; Yu & Smith, 2007), or by relying on knowledge that each word in a language tends to only have one meaning ("mutual-exclusivity"; Lewis, Cristiano, Lake, Kwan, & Frank, 2020; Clark, 1987; Markman & Wachtel, 1988). While these various learning strategies, and others, could in principle be applied to verbs, there is reason to think that these strategies may be better suited for concrete nouns, where the perceptual information is more stable and less complex. Consistent with this prediction, verbs are typically learned later in development than concrete nouns (e.g. Bates et al., 1994; Frank, Braginsky, Marchman, & Yurovsky, in press). Understanding the relative strength of various learning strategies for mapping word forms to word meanings could shed light on the source of the developmental lag in verb learning.

In addition to uncertainty about the overall strength of the syntactic bootstrapping effect, there are a number of open theoretical questions about the nature of the effect. First, how does the strength of the effect change across development? One possibility is that the ability to use syntactic bootstrapping to learn novel verbs is unlearned or available early on in development as soon as the relevant syntactic information is represented, and the ability does not strengthen with development (Fisher, Jin, & Scott, 2020). An alternative possibility is that the effect becomes stronger with maturation and experience. There are a range of reasons to think that the effect might strengthen. For example, with development, children's vocabulary size increases, making it more likely they will know the nouns in the linguistic input containing a novel verb. If the effect becomes stronger with development, this would suggest that syntactic bootstrapping might become an increasingly powerful learning mechanism for children as they learn more words. [learning to learn stuff] Prior work has examined the syntactic bootstrapping effect across different age groups (e.g. Jin, 2015; Gertner & Fisher, 2012) but it is unclear what the developmental trajectory of this effect looks like. Characterizing this developmental change is important to understanding the role that different word learning strategies play across stages of language acquistion.

Second, how robust is the effect to different syntactic structures? Syntactic bootstrapping is theorized to be a general learning mechanism that could in principle be applied to the range of syntactic structures in children's input. However, some experiments have found a syntactic bootstrapping effect for transitive sentences, but not intransitive sentences (Arunachalam & Waxman, 2010; Yuan, Fisher, & Snedeker, 2012), whereas others find the opposite pattern (Bunger & Lidz, 2004; Naigles & Kako, 1993). Plausible explanations for both patterns have been proposed. For example, the effect may be stronger for transitive sentences relative to intransitives, because intransitive sentences are relatively more flexible in their usage. Consider, for instance, the sentence "the girl is daxing." This sentence could describe either a scene in which a girl is doing a transitive action (e.g., patting a boy) or a scene in which the girl is doing an intransitive action (e.g., jumping). In contrast, the sentence "the girl is daxing the boy" could only describe a scene in which the girl is doing a transitive action (e.g., patting the boy). This asymmetry may lead children to treat the syntactic information in sentences with a transitive novel verb as a stronger cue to a verbs's meaning, relative to sentences with an intransitive novel verb. On the other hand, intransitive sentences have fewer processing demands relative to transitive nouns (Lidz, Bunger, Leddon, Baier, & Waxman, 2009), thereby potentially making syntactic bootstrapping easier for intransitive nouns.

The literature has also revealed conflicting findings about the robustness of syntactic bootstrapping to different types of noun phrase structures. In everyday discourse, people and objects are often referred to by pronouns (e.g., "she") instead of descriptive nouns ("the girl"). There is some evidence in the literature that semantically rich, descriptive nouns are beneficial for verb learning because the semantic content of the surrounding nouns can help scaffold the interpretation of the verbs (Arunachalam & Waxman, 2015; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005). Others, however, have found that pronouns reduce processing load, thereby facilitating syntactic bootstrapping (Childers & Tomasello, 2001; Lidz et al., 2009). In the case of both predicate types and noun phrase structures, it is difficult to evaluate the robustness of syntactic bootstrapping to variability in the linguistic input without a clear understanding of the empirical pattern.

One challenge in addressing these theoretical questions is the range of methodological implementations of tests of the syntactic bootstrapping phenomenon across the the literature. The presence of methodological variability means that it is not clear if observed differences in outcomes are due to methodologial variability, sampling error, or theoretical moderators of interests. For example, across studies, there is variability in the relative onset of the linguistic and visual stimuli. In some studies, children see the events at the same time or soon after hearing the relevant sentences (Gertner & Fisher, 2012; Naigles, 1990), while in others, the sentences are accompanied by an irrelevant scene (e.g. a person on the phone talking or two people conversing) followed by the target visual stimuli (Arunachalam & Waxman, 2010; Yuan & Fisher, 2009). The lag between the linguistic stimuli and the visual stimuli likely increases the memory demands of the task, and could influence children's ability to identify the correct referent. Critically, if this methodological difference co-varies with the age of children tested, it may be difficult to draw conclusions about the developmental trajectory of the effect across studies.

The meta-analytic method provides a powerful analytic tool for quantifying theoretically important effects, and has increasingly been applied in the language acquisition literature (Bergmann et al., 2018; Cristia, 2018; Lewis et al., 2020; Rabagliati, Ferguson, & Lew-Williams, 2019). Meta-analysis involves quantifying the size of an effect in individual experiments with a standardized measure (such as Cohen's d), and then aggregating across studies statistically to estimate an overall effect size. Because a meta-analysis reflects estimates from many more participants than any individual study, the meta-analytic method has greater power to detect a true effect and precisely quantify it. The meta-analytic method also allows researchers to explore moderators of an effect. Detecting a moderator to a small effect requires large sample sizes, but infant experiments, like those investigating syntactic bootstrapping, typically have small sample sizes (Bergmann et al., 2018; Oakes, 2017). These small sample sizes mean that individual studies have low power to detect a true moderating effect, leading to literature with many null and conflicting effects for moderating influences. Meta-analysis addresses this limitation by providing a higher-powered test of theoretically important factors, as well as revealing variability in effect sizes due to methodological variability across studies.

The plan for the paper is as follows. We first describe our method for conducting a meta-analysis of the syntactic bootstrapping effect. We then assess the evidential value of this literature by evaluating the influence of publication bias. Next, we characterize the size of the effect with respect to other word learning phenomena. We then examine the moderating influence of experience, predicate type, and noun phrase structure on the syntactic bootstrapping effect, and the relationship between theoretical moderators and methodological variability in our meta-analytic data. We find some evidence for publication bias in the syntactic bootstrapping literature, but that there is nevertheless a small syntactic bootstrapping effect. The effect is larger for transitive sentences, relative to intransitive sentences, but does not appear to be influenced by experience, noun phrase structure, and a range of methodological factors. In the General Discussion, we discuss the implications of our findings for theories of verb learning.

Method

Literature Search

We conducted a literature search of the syntactic bootstrapping literature following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist (PRISMA; Moher, Liberati, Tetzlaff, Altman, & Prisma Group, 2009). We identified relevant papers through a keyword search in Google Scholar with the phrase "syntactic bootstrapping" and a forward search on papers that cited the seminal paper, Naigles (1990) (total records identified: N = 3,339; retrieved between May 2020 and July 2020; Figure 1). We screened the abstracts for relevance of the first 60 pages of the keyword search results (N = 600) and the first 10 pages of the forward search results (N = 100). The screening processes ended because we could no longer identify relevant, non-duplicate papers from consecutive pages. Additional papers were identified by consulting the references section of the most recent literature review (N = 155; Fisher et al., 2020) and experts in the field (N = 11). Our sample included published journal articles, conference proceedings, doctoral dissertations, and unpublished manuscripts. We refer to these collectively as "papers" in the following sections.

We restricted our final sample to papers that satisfied the following criteria: First, the experimental paradigm involved a two-alternative forced-choice task in which participants were instructed to identify the scene that matched the linguistic stimuli. Second, the visual stimuli were two events displayed side-by-side on a computer monitor. One event depicted a causative action (e.g., one agent causes the other to move), and the other a non-causative action (e.g. two agents move simultaneously but do not causally interact with each other). We included studies with either videos of live actors or animated clips. Third, the linguistic stimuli included at least one novel verb embedded in a syntactically informative frame. For example, "Look, it's kradding!" embeds the novel verb in an intransitive syntactic frame that is informative about the meaning of the novel verb "kradding". In contrast, "Look,

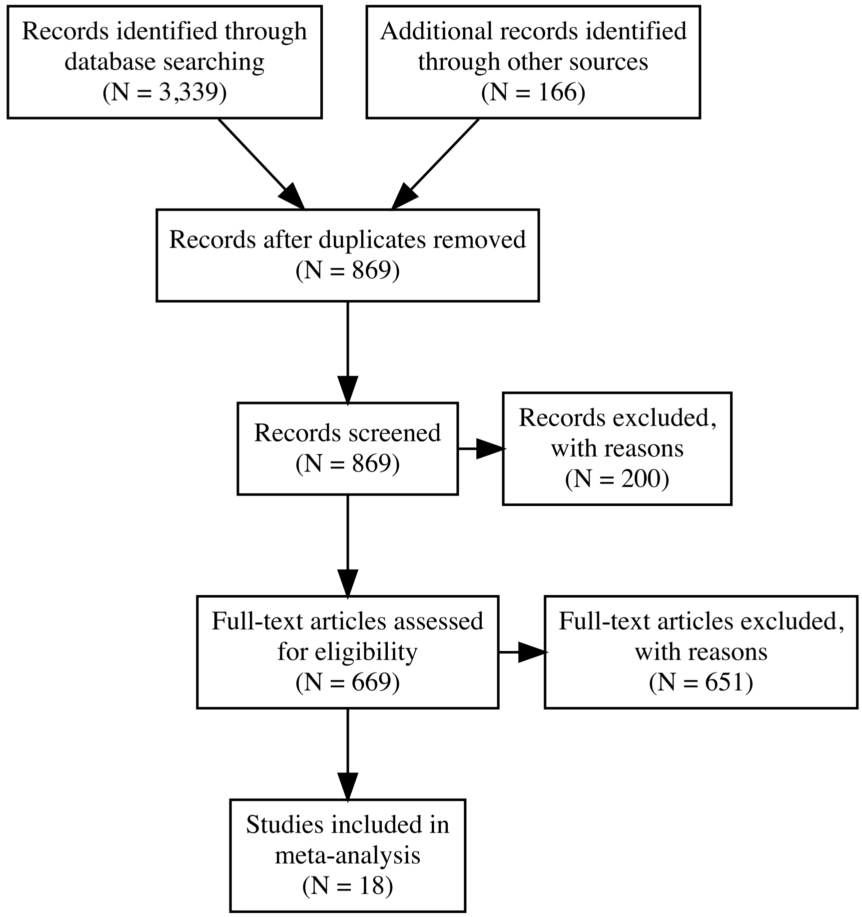

Figure 1. PRISMA plot showing literature review process. Numbers indicate papers at each stage of the review review process. Our meta-analysis included a final sample of 18 papers.

kradding!" does not provide informative syntactic information. Finally, we restricted our sample to studies with English-speaking, typically-developing children. Papers that satisfied these constraints reflected a range of methodological implementations which we examine systematically below (see Moderators section below). Papers included in our meta-analysis are marked with an asterisks in the references.

Our final sample included 18 papers, corresponding to 60 individual effect sizes and 730 unique infants (mean age: 24 months; 28 days).

Data Entry

For each paper, we entered metadata about the paper (e.g., citation), information to calculate effect sizes, and moderators. We entered a separate effect size for each experimental manipulation and age group per paper ("conditions"). Most papers therefore contained multiple conditions, corresponding to multiple effect sizes in our meta-analysis.

Calculating individual effect sizes. For each condition, we recorded the sample size, the mean proportion correct responses, and the across-participant standard deviation of proportion correct responses. The mean and standard deviation were obtained from one of the four sources: a) the results section or the data-presenting tables (N = 37); b) the plots (N =10); c) the original authors (N = 12); d) imputation using values from studies with similar designs (N = 1, Hirsh-Pasek, Golinkoff, & Naigles (1996); the missing SD values were imputed from Naigles (1990)). Previous work suggests using imputed values from highly similiar studies improves the accuracy of effect size estimates (Furukawa, Barbui, Cipriani, Brambilla, & Watanabe, 2006). The reported results do not qualitatively change when conditions from Hirsh-Pasek et al. (1996) are excluded from our sample (see SI).

Using the raw coded data, we calculated an effect size estimate for each condition as Cohen's d. Cohen's d was calculated as the difference between the proportion correct responses and chance (.5), divided by a pooled estimate of variance (see SI for example calculation). Note that we assume baseline performance to be .5 in all cases, even when an empirical baseline was reported (N = 6 conditions). This analytical decision was made in order to standardize the effect size estimate across all conditions in our sample, most of which did not report an empirical baseline.

Moderators. For each effect size in our sample, we coded several theoretical and methodological variables. The information was either retrieved from the methods section of the paper or by contacting authors.

Four theoretical variables were coded: participant age, participant vocabulary size, predicate type, and noun phrase type. Participant age was entered in mean age in months ($N = 60$). Vocabulary size was recorded as the median productive vocabulary measured by MacArthur-Bates Communicative Development Inventories (CDI) Words and Sentences (Fenson, 2000; $N = 32$). Predicate type was coded as either transitive ($N = 30$) or intransitive ($N = 30$). Noun phrase type encoded information about the agent verb argument of the sentence stimulus. The agent of the sentence was coded as being either noun (e.g., "the girl"; $N = 26$) or pronoun ("she"; $N = 34$). A condition was coded as "pronoun" if it contained at least one instance of a pronoun referring to the agent.

In addition to the theoretical variables, we coded a range of methodological variables that varied across the studies in our sample. First, we coded whether the paradigm included a practice trial prior to the testing phase. A study was coded as having a practice trial if there was at least one trial in which children were presented with a familiar verb and asked to identify a familiar action (e.g. "Find jumping"; N with practice phase = 36). Second, we coded whether or not the paradigm involved trials in which children were prompted to identify the nouns in the testing events (e.g., "Where's the bunny?"; N with character identification phase = 16). Third, we coded whether the linguistic and visual stimuli were presented synchronously with each other ("Stimuli Synchronicity"). An experimental condition was coded as "asynchronous" if the linguistic stimulus was first paired with an irrelevant visual scene (e.g. a person on the phone talking), and the matching visual stimulus was not shown until the training phase is over ($N = 37$); a condition was coded as "simultaneous" if the very first training sentence was presented along with the visual stimuli depicting the relevant action or along with an attention-getter or a blank screen, immediately followed by the relevant action ($N = 23$). The synchronicity between the linguistic stimuli and the visual stimuli could potentially scaffold the mapping process. Fourth, we coded the temporal distribution of the training and the testing trials (Mass: $N = 28$; Distributed: $N = 32$). A procedure was categorized as "mass" if participants were trained exclusively on one novel verb and tested on the same verb, and "distributed" if they were trained and tested on multiple novel verbs. The temporal distribution is linked to the amount of learning experience children have prior to the test. Finally, we coded how many times each novel verb was spoken in a syntactically-informative way during training.[1]

Analytic Approach

We analyzed the data using multi-level random effect models implemented in the metafor package in R (Viechtbauer, 2010). The random effect structure included groupings by paper and participant group in order to account for the clustering of effect sizes in our sample. The sensitivity analysis was conducted using the PublicationBias package in R (Mathur & VanderWeele, 2020). Moderator variables were included as additive fixed effects. All estimate ranges correspond to 95% confidence intervals, unless otherwise noted. Data and analysis scripts are available at XXXX, and the dataset can be interactively explored at XXXX.

Results

Figure **??** shows effect size estimates for all conditions in our sample. The weighted mean effect size was 0.22 [0.01, 0.43], which significantly differed from 0 (Z = 2.06; p = 0.04). There was evidence for considerable heterogenity in effect sizes across our sample (Q = 196.07; p <.001), meaning that there is unexplained variance in effect sizes aross studies.

Evidential value of the syntactic bootstrapping literature

We first evaluated the evidential value of the literature by assessing the evidence for publication bias. The intuition underlying these analyses is that, due to random variation, a literature should be expected to contain studies both with and without statistically

[1] See SI for additional methodological moderators. These additional moderators overlap substantially with the target moderators of interest presented in the Main Text.

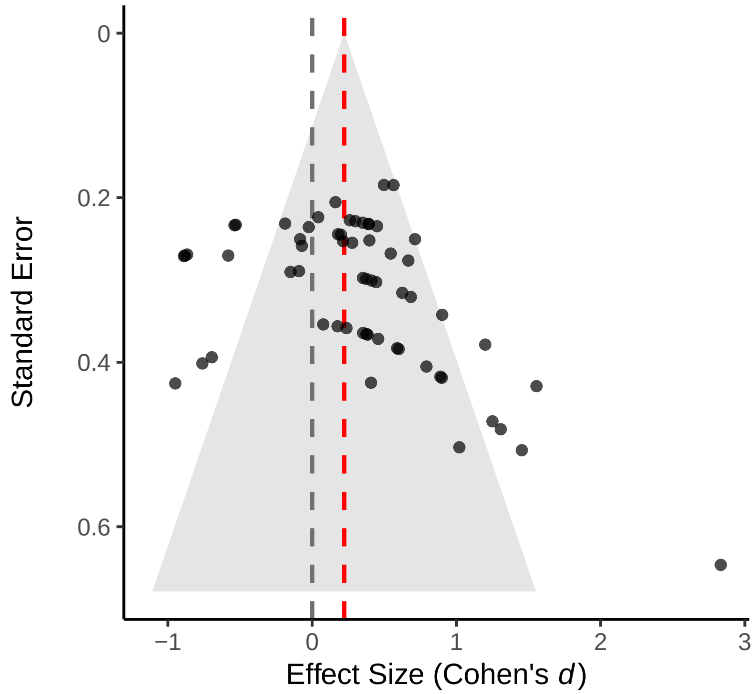

Figure 2. Funnel plot showing the standard error of each effect size estimate in our meta-analysis as a function of the magnitude of that effect size. The gray and red vertical dashed lines correspond to an effect size of zero and the meta-analytic effect size estimate, respectively. The grey funnel represents a 95% confidence interval around meta-analytic estimate. In the absence of publication bias, effect size estimates should be symmetrically distributed around the red line.

significant effect sizes for the target phenomenon. Critically, however, publication pressures may lead researchers to be more likely to publish findings with statistically significant results, resulting in a biased literature. The absence of these studies from the meta-analysis yields a meta-analytic estimate that over-estimates the true effect size, and threatens the evidential value of the literature. We present two analyses that assess publication bias in the syntatic bootstrapping literature: a classic funnel plot analysis, and a sensitivity analysis that assumes a more plausible model of the publication process.

Figure 2 presents the funnel plot for the effect sizes in our sample. A funnel plot shows estimates of effect size variance (plotted with large values lower on the axis) as a function of the magnitude of the effect size (Egger, Smith, Schneider, & Minder, 1997). Under a model of publication bias in which researchers decide whether or not to publish a study based on the magnitude of its effect size (larger effect sizes being more likely), effect size estimates should fall symmetrically around the grand effect size estimate. Evidence of asymmetry around the grand mean, particularly more large, positive effect sizes, would suggest that the literature reflects a biased sample of studies. A formal test of asymmetry in our sample revealed evidence for asymmetry (Egger's test: Z = -3.55 and 4.62; $p <$ .0001).

The funnel plot analysis provides some evidence for publication bias, but the interpretation of this analysis is limited by the fact that it assumes a relatively implausible model of how researchers decide which studies to make public: the criteria for publishing a study in a journal is typically not the *size* of the effect, as assumed by the funnel plot analysis, but rather whether or not the p-value of the hypothesis test for that effect is below some threshold (usually .05). We therefore conducted a second analysis of publication bias, called a sensitivity analysis (Mathur & VanderWeele, 2020), which assumes that the decision to publish results is determined by the size of the p-value, rather than the magnitude of the effect size.

The goal of the sensitivity analysis is to determine how sensitive the meta-analytic effect size is to "missing" non-significant studies. Critically, because the degree of publication is not known (i.e. the degree to which significant results are more likely to be published, relative to insignificant results), the sensitivity analysis assumes a worst-case publication bias scenario and estimates the meta-analytic effect size under this scenario. The worst-case scenario assumed by the model is that significant studies are infinitely more likely to be published than non-significant studies.[2] A meta-analytic effect size under this scenario can be estimated by analyzing only those studies with significant effect size estimates.

[2] Technically, the model assumes studies with effect sizes that are statistically significant ($p < .05$) *and* greater than zero are infintely more likey to be published. See Mathur and VanderWeele (2020) for additional details.

Conducting this sensitivity analysis on our data reveals that no amount of publication bias could attentuate the point estimate of the effect size to 0. Nevertheless, the worst-case scenario appreciably attenuates the meta-analytic effect size, and the attenuated effect size estimate includes 0 in its 95% confidence interval (0.08 [-0.1, 0.26]; see SI for additional details).

In sum, across two types of analyses, we find some evidence for publication bias in the syntactic bootstrapping literature, but even under worst-case scenarios publication bias was not enough to fully attenuate the meta-analytic point estimate to 0. Further, some of the publication bias observed in the funnel plot analysis may be due to heterogeneity in the data. In the following sections, we analyze theoretical and methodological moderators that may contribute to this heterogeneity, though we emphasize that the likely presence of publication bias implies that these moderators should be interpreted with caution.

Relating the syntactic bootstrapping effect to other word learning strategies

How does the strength of the syntactic bootstrapping effect compare to that of other word learning strategies? To answer this question, we compared the meta-analytic syntactic bootstrapping effect size to effect sizes for other word learning strategies estimated from a meta-analysis of each literature. We considered an opportunity sample of word learning strategies, based on those strategies with available meta-analytic data. In particular, we selected all word learning strategies available in a database of language acquistion meta-analyses, called Metalab (Bergmann et al., 2018). We included a word learning strategy in our analysis if it could be considered to facilitate an inference about which word form was associated with a particular object. This allowed for the comparision of the syntactic boostrapping effect to four additional word learning strategies: (i) mutual exclusivity (Clark, 1987; Lewis et al., 2020; Markman & Wachtel, 1988), assuming a novel word refers to a novel object, (ii) cross-situational word learning (Yu & Smith, 2007), tracking word-object co-occurrences across situations, (iii) gaze following (Frank, Lewis, &

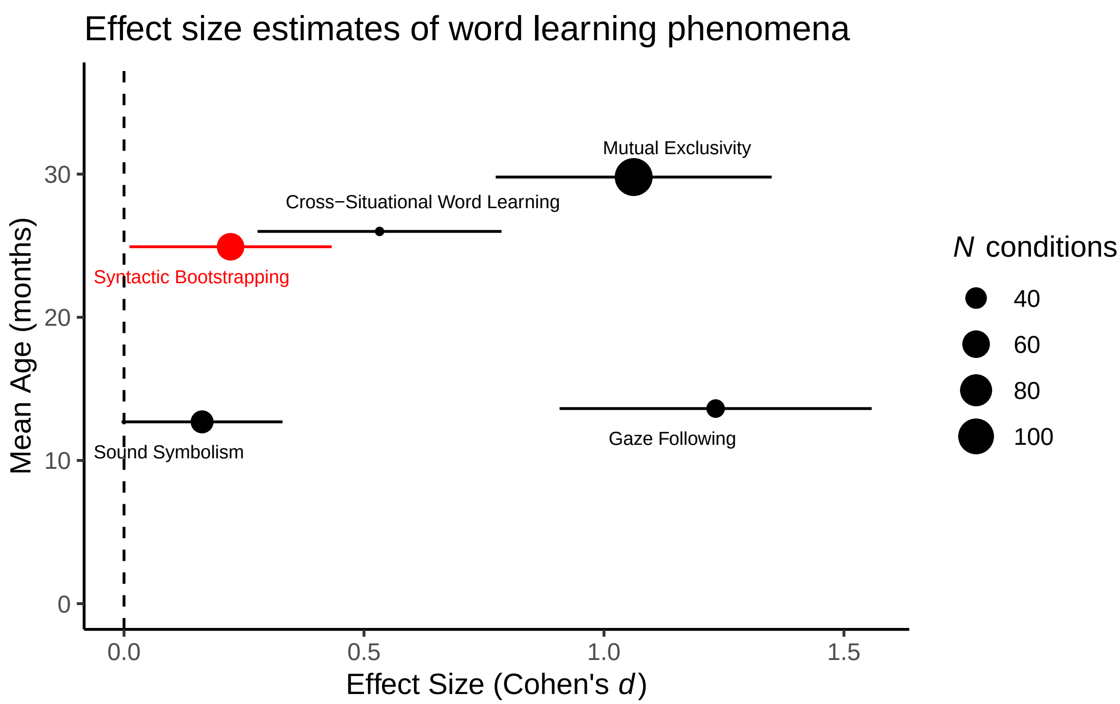

Figure 3. Meta-analytic effect sizes of five word learning phenomena, including syntactic bootstrapping (red). Point size corresponds to the number of individual conditions included in each meta-analysis. The x-axis shows the magnitude of the meta-analytic effect size estimate; the y-axis shows the mean age in months of children in each meta-analysis.

MacDonald, 2016; Scaife & Bruner, 1975), following the eye gaze of a speaker to intended referent, and (iv) sound symbolism (Fort et al., 2018), exploiting sound-meaning regularities in the lexicon. While these four strategies are not exhaustive of the strategies that have been proposed, they are representative of the major theoretical perspectives in the word learning literature, including constraints and biases (Markman, 1990), statistical learning (Romberg & Saffran, 2010), and communicative inferences (Tomasello, 2010). For each of these four comparison strategies, we calculated the meta-analytic effect size using the same model specification as for syntactic bootstrapping, restricting the sample to studies with a mean age of children younger than 48-month-olds.

Figure 3 shows the meta-analytic effect size for syntactic bootstrapping and each of the other four word learning stategies. The syntactic bootstrapping effect size (0.22 [0.01, 0.43]) was comparable in size to that of sound symbolism (d = 0.16 [-0.01, 0.33]; Mean age:

386.43 Days) and cross-situational learning (d = 0.53 [0.28, 0.79]; Mean age: 791.38 Days), and less than a quarter of the size of both mutual exclusivity (d = 1.06 [0.77, 1.35]; Mean age: 906.97 days) and gaze following (d = 1.23 [0.91, 1.56]; Mean age: 414.73 Days). Importantly, the small effect size of syntactic bootstrapping relative to mutual exclusivity and gaze following cannot be due alone to differences in the ages of the samples in these different meta-analyses, since participants in the syntactic bootstrapping meta-analysis were older on average than those in the gaze following meta-analysis, and roughly the same age as those in the mutual exclusivity meta-analysis.

Theoretical Moderators

We next asked whether the overall effect size estimate was moderated by our theoretical moderators of interest: development-related moderators (vocabulary and age), and sentence structure moderators (predicate and noun phrase types).

Development. How does the strength of the syntactic bootstrapping effect change across development? We examined two measures of developmental change: age (months) and vocabulary size. These two measures were strongly correlated with each other ($r(30)$ = 0.85, p <.0001). There was no effect of age (β = -0.02 [-0.03, <.001], SE = 0.01, z = -1.64, p = 0.1; Fig. 4) or vocabulary (β = -0.01 [-0.02, <.001], SE <.001, z = -1.93, p = 0.05) on effect size (Fig. 5a).

Sentence structure. We next asked how properties of the sentence struture that children heard influenced the strength of the syntactic bootstrapping effect. Predicate type (transitive vs. intransitive) was a significant moderator, (β = 0.24 [0.02, 0.46], SE = 0.11, z = 2.13, p = 0.03), with the effect being larger for transitive conditions (M = 0.49, SD = 0.72) relative to intransitive conditions (M = 0.17, SD = 0.58). In contrast, there was no effect of agent argument type (pronoun vs. noun; β = 0.14 [-0.26, 0.53], SE = 0.2, z = 0.69, p = 0.49).

To compare the effects of all theoretical moderators, we fit an additive model with all

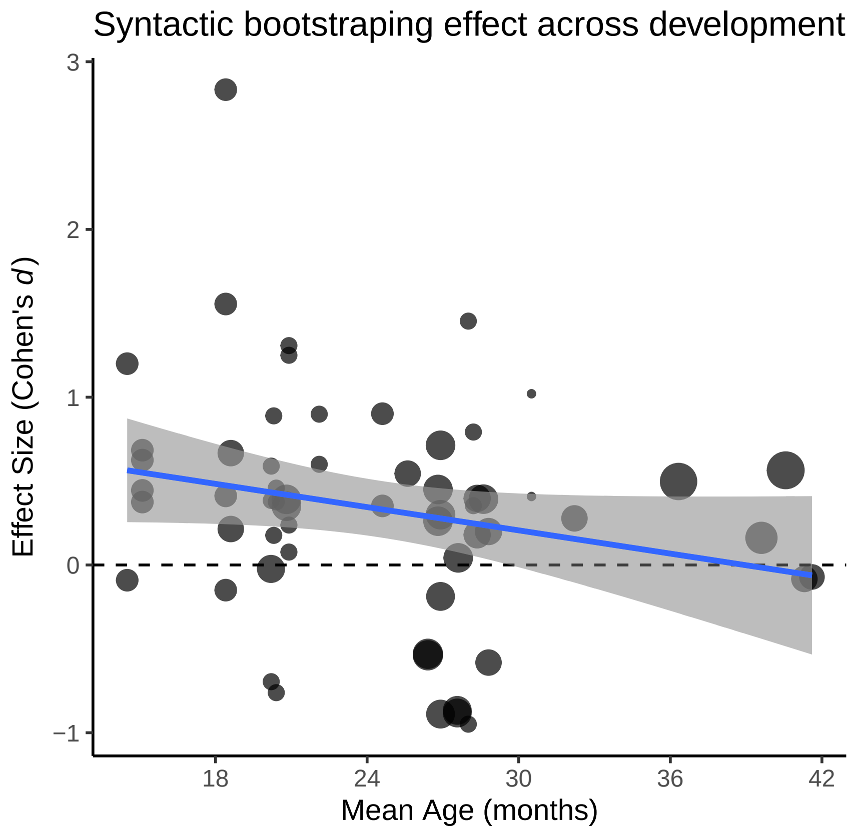

Figure 4. Syntactic bootstrapping effect size (Cohen's d) as a function of age in months. Each point corresponds to one effect size (condition), and point size corresponds to the number of children in that condition. The blue line shows a linear model fit and the corresponding standard error. The slope of the model fit does not significantly differ from zero, suggesting no appreciable developmental change in the size of the syntactic bootstrapping effect as a function of age. The dashed line indicates an effect size of zero.

theoretical variables as fixed effects. We excluded vocabulary size from the additive model since it was highly correlated with age, and only available for a subset of conditions ($N =$ 32). Figure 5 shows estimates for each of the single-predictor models along with the additive linear model. The additive model revelaed estimates that were highly comparable to the single-predictor model.

In summary, we find that predicate type is a significant predictor of the effect size: conditions with transitive sentences were associated with larger effect sizes than those tested with intransitive sentences. No other theoretical variable significantly moderated the syntactic bootstrapping effect.

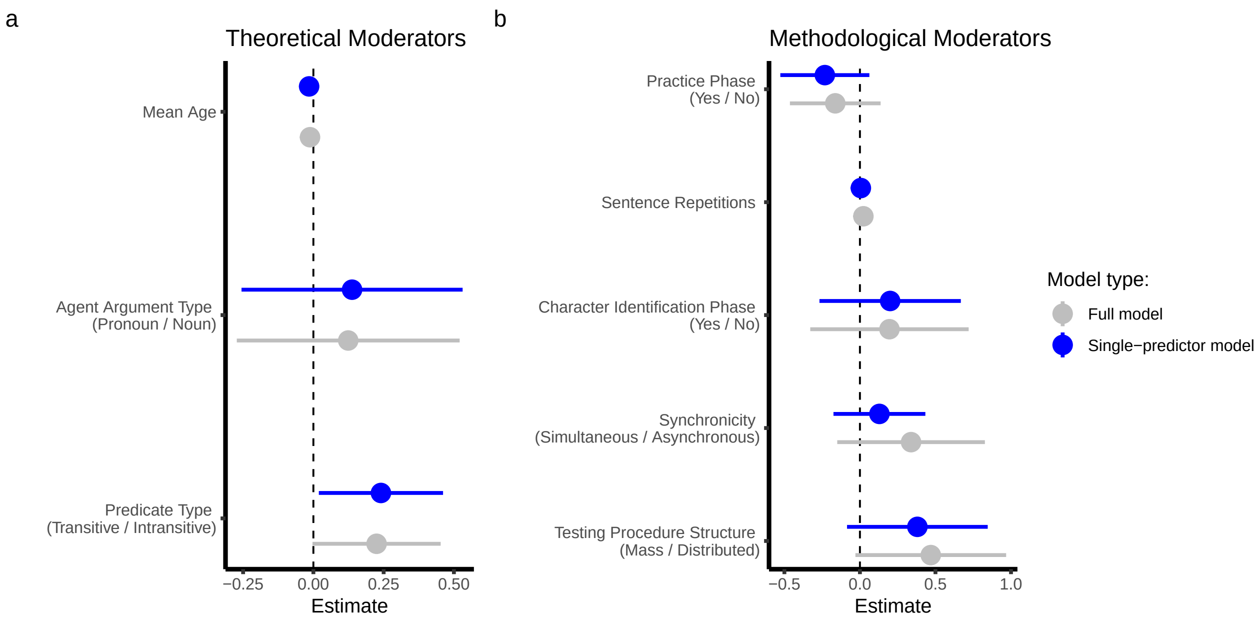

Figure 5. Meta-analytic models parameter estimates for (a) theoretical and (b) methodological moderators. Blue points show model estimates from single-predictor model; grey points show model estimates from additive linear model with all moderators included. Ranges correspond to 95% confidence intervals. Levels for categorical variables are given in parentheses, with the first level indicating the base level in the model.

Methodological Moderators

One limiting factor in interpreting the moderating role of theoretical variables is that there was appreciable variability across studies in the exact method used in testing children. It is possible that this methodological variability conceals true underlying moderating influences. For example, if researchers adapt their method to the age of the children they are targeting, developmental change in the strength of the effect may not be detectable (Bergmann et al., 2018).

To evaluate this possiblity, we asked whether the five different methodological variables (practice phase, sentence repetitions, character identifcation phase, synchronicity, testing procedure structure) moderated the syntactic bootstrapping effect. None of these methodological variables were significant moderators of the effect in a single predictor model (Fig. 5; see SI for exact estimates). In an additive linear model with all five methodological predictors, there was a marginally significant effect of testing procedure structure ($\beta = 0.47$ [-0.03, 0.97], $SE = 0.25$, $z = 1.84$, $p = 0.07$), with mass testing designs tending to have larger effect sizes than distributed designs. This finding suggests that children tested in a procedure with only one train-test pair performed better than those tested in procedure with multiple train-test pairs. Finally, we asked how these methodological moderators related to our theoretical moderators of interest. Controlling for methodological variables did not qualitiatively change the role of any of the theoretical moderators (see SI). Taken together, these analyses suggest that methodological variables do not play a large influencing role on the size of the syntactic bootstrapping effect.

General Discussion

- interpret with caution - small es, and there's publication bias!
- Explanations for lack of age effect: variability in design as a function of age, change in input, etc. [Meta point: measuring developmental change requires constancy in task]
- Transitivity effect is robust - implications for this? Cross-linguistic implications? Text from old intro about multiple possible explanations of this

Three decades of research in syntactic bootstrapping has tremedously extended our understanding of early verb acquisition. Here we built upon the previous literature and presented a quantitative synthesis. We aggregating the results from [] unique infants from [] experimental conditions, and we found that despite some evidence for publication bias, there is a small effect size (). We found no effect of age, but there is a weak negative effect of productive vocabulary size. Furthermore, the predicate type is a robust predictor: using transitive sentences at test postively predict the effect size. No methodological moderators considered were significant, suggesting that the myriad implementational differences do not influence the learning outcome. In the following secionts, we considered the implications of our findings.

Small effect size

The effect size for syntactic bootstrapping is considered small by the conventional Cohen's d interpretation (STANDARD). Relatively, it is also small compared to other word learning strategies. We compared syntactic bootstrapping with an opportunity sample of early language development meta-analyses: mutual-exclusivity, cross-situational learning, gaze following and sound symbolism. These strategies and biases are selected because they are considered to be instrumental for young children to establish the mapping relationships between words and their references in a local context. We showed that the effect size for Syntactic Bootstrapping is comparable to sound symbolism but smaller than the all other three phenomena.

One possible explanation for this discrepancy is the syntactic category difference. Studies in the four domains are all primarily concerned with noun learning. However, syntactic bootstrapping deviates by expecting infants to learn verb. The challenging nature of verb learning relative to noun learning has been well documented in the observational literature (CITE) and experimental studies (CITE). Second, studies in syntactic bootstrapping may be more challenging for children because they require children to attend, process and incoportate the full sentence. The neighboring words and their orderings are essential to determine the references of the novel word (e.g. "The bunny is *gorping* the duck" and "the duck is *gorping* the bunny" corresponds to two distinct actions). In contrast, in a classic mutual exclusivity study, the local linguistic context does not play a central role in the semantics of the novel word (e.g. "Look at the *dax*" can be replaced by "Can you show me the *dax*").

Lack of age effect

We also find an intriguing lack of age effect. Children do not seem to get better at this task as they get older, which is also in contrast with other phenomena (MAYBE LIST OUT COEFFICIENTS?). One potential cause for the lack of age effect is the nature of this learning strategy: syntactic bootstrapping has been argued to be an unlearned bias (CITE), and therefore the accumulation of experience should not have a significant impact on this effect as children age. Of coruse, an unlearned bias does not necessarily imply that the effect sizes should not change. In theory, older children tend to process information faster, have more capacity in their working memory, and are able to stay focus at test longer (CITATION?). As a result, another possible cause for the lack of age effect is the way researchers adapting studies based on children's age. Although in our studies we did not find adding methodological moderators to significantly change the age model's results, it is plausible that our coded moderators fail to capture the changes the methods. Indeed, an intuitive inspection of the visual stimuli use change as a function of the age suggests suggests that researchers tend to use more complex visua stimuli to older children (see SI). With older infants tested in a more demanding way, the adaptation could mask some developmental changes of the learning strategies. [FIT MODEL IN THE SUBSET?]

```
##
## Multivariate Meta-Analysis Model (k = 32; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -30.8670   61.7339   71.7339   78.7399   74.2339
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed                           factor
## sigma^2.1  0.1411  0.3757      5     no                       short_cite
## sigma^2.2  0.1010  0.3178     32     no            short_cite/same_infant
## sigma^2.3  0.1010  0.3178     32     no  short_cite/same_infant/row_id
##
## Test for Residual Heterogeneity:
```

```
## QE(df = 30) = 103.6121, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 3.7252, p-val = 0.0536
##
## Model Results:
##
##                          estimate      se     zval    pval    ci.lb   ci.ub
## intrcpt                    0.5869  0.2983   1.9672  0.0492   0.0022  1.1717  *
## productive_vocab_median   -0.0086  0.0045  -1.9301  0.0536  -0.0174  0.0001  .
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Multivariate Meta-Analysis Model (k = 32; method: REML)
##
##   logLik  Deviance       AIC       BIC      AICc
## -31.4570   62.9140   72.9140   79.9200   75.4140
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed                          factor
## sigma^2.1  0.2079  0.4559      5     no                      short_cite
## sigma^2.2  0.1005  0.3170     32     no          short_cite/same_infant
## sigma^2.3  0.1005  0.3170     32     no  short_cite/same_infant/row_id
##
## Test for Residual Heterogeneity:
```

```
## QE(df = 30) = 119.9549, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 2.2128, p-val = 0.1369
##
## Model Results:
##
##           estimate      se     zval    pval    ci.lb   ci.ub
## intrcpt     0.8353  0.5076   1.6456  0.0998  -0.1596  1.8301  .
## mean_age   -0.0009  0.0006  -1.4876  0.1369  -0.0020  0.0003
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Implications of the robust transitivity effect

Children perform better in conditions using transitive sentences than those using intransitive sentences. This finding consolidates the observations many researchers have made [CITATION]. Three factors may have given rise of this transitivity advantage. First, under certain contexts, the transitive sentences have less ambiguity than their intransitive counterparts. If one sees two scenes: in one a girl patting a boy, in the other a girl and a boy jumping side-by-side, then upon hearing “the girl is gorping the boy”, only the scene with patting action is a plausible interpretation. In contrast, if one hears the intransitive sentence, “the girl is gorping”, the verb can be interpreted as being consistent with both the causative scene and the non-causative scene. Second, children may have more experience with transitive sentences than intransitive sentences. A corpus analysis on parental utterances for 1- to 6-year-old children found that transitive sentences make up approximately 24.36% of all utterances, whereas intransitive sentences occupy only 17.24%

(Laakso & Smith, 2007). Thus the familiarity with transitive sentences can potentially improve children's understanding of transitive sentences in the lab. Last but not least, the perceptual saliency of the visual scenes may have contributed to the asymmetrical performance in the transitive conditions and the transitive conditions. Children were found to have a baseline preference for two actors event over one actor event (e.g., Yuan et al., 2012), and for synchronous movement over causative movement (e.g., Naigles & Kako, 1993). These preferences can make detecting an effect in the transitive condition easier than in the intransitive condition.

Limitations and moving forward

Children build up their vocabularies at a dazzling speed. Toward the end of the second year, they were estimated to have a productive vocabulary size of around 300 (Fenson et al., 1994). But children's impressive word-learning abilities do not apply equally to words of all kinds. Verbs, for example, constitute a special challenge for children. Numerous observational studies have shown that verbs are often learned later than nouns (Goldin-Meadow, Seligman, & Gelman, 1976; Lieven, Pine, & Barnes, 1992; Longobardi, Spataro, Putnick, & Bornstein, 2017; Nelson, 1973; Nelson, Hampson, & Shaw, 1993), and laboratory experiments also found that children's verb learning is more slowly and difficult than noun learning(Abbot-Smith, Imai, Durrant, & Nurmsoo, 2017; Childers, Heard, Ring, Pai, & Sallquist, 2012; Childers & Tomasello, 2002, 2006; Gentner, 1978; Imai et al., 2008; Oviatt, 1980; Schwartz & Leonard, 1984). Some scholars have argued that the challenge of verbs is language-dependent (Choi & Gopnik, 1995; Tardif, 1996; Tardif, Gelman, & Xu, 1999). Nevertheless, the cross-linguistic findings are mixed, with others found consistent patterns across languages (Au, Dapretto, & Song, 1994; Bornstein et al., 2004; Kim, McGregor, & Thompson, 2000; Papaeliou & Rescorla, 2011). A more recent large scale, cross-linguistic corpus analysis has shown that across 23 different languages, verbs and other predicates, when compared with nouns, constitute a smaller proportion of children's early production vocabularies, though there are more cross-linguistic differences in early comprehension vocabularies (Frank et al., in press). Together, these findings suggest that at least certain aspects of the challenge associated with verb learning are universal across languages. Some have hypothesized that it is due to the nature of verbs' references: Unlike nouns, the references of verbs are ephemeral and ever-changing. As a consequence, to learn verbs, children need to rely less on the unreliable extralinguistic information and rely more on the information within the verbs' linguistic context, such as syntax(Gentner & Boroditsky, 2001, p. Gentner (2006)).

The idea that syntactic information can shape children's interpretations of novel words was not new. As early as the late 1950s, a classic study by Brown (1957) has shown that 3- to 5-year-old children would choose to map a novel word to an action if the word is introduced in a verb context (e.g."Do you know what it means to sib?"), and map to an object if introduced in a noun context (e.g."Do you know what a sib is?"). However, this early finding only provided evidence for older children's abilities to discriminate between syntactic categories. It remained unknown how early the abilities emerge and how detailed the underlying representations for the verbs are. In the late 1980s and early 1990s, a hypothesis known as "Syntactic Bootstrapping" revived the interests in the early interaction between syntax and verbs' semantics during language development (Gleitman, 1990; Landau & Gleitman, 1985). The basic idea of Syntactic Bootstrapping is that the structures of the sentences were potent cues for children to interpret the meanings of the verbs. This idea soon received empirical support from a seminal study by Naigles (1990). By using the Intermodal Preferential Looking Paradigm (Golinkoff et al., 1987), Naigles (1990) provided the first evidence for 25-month-olds children's abilities to incorporate syntactic cues in verbs learning.

In this seminal study, the children were tested on four different novel verbs. For each novel verb, there was a training phase and a testing phase. During the training phase, the children would hear either a series of transitive sentences or intransitive sentences, depending on the conditions they were assigned to. Each sentence grammatically contained the novel verb (e.g., for transitive: “The duck is gorping the bunny”; for intransitive: “The duck and the bunny are gorping”.) The children were also exposed to the visual stimuli depicting a two-actor action while listening to these sentences. The two actors, one in a bunny suit and the other in a duck suit, would perform two novel actions simultaneously. They would move their arms synchronously while the one pushed the other bending forward. After the training phase, the children would be prompted to look for the action (e.g. “Where’s the gorping now?”). The children faced two screens side-by-side, one would show the bunny and duck performing non-causal synchronous arm movements, while the other would show the causal pushing-and-bending action. The children’s fixation time toward each of the screens was recorded. A longer looking time for either of the screen indicated a match between the children’s conceptual representations of the novel verbs and the action on the screens. Naigles (1990) found that children assigned to the transitive condition looked longer at the causal action than the non-causal action, whereas the children in the intransitive condition showed the opposite pattern. This indicated that at this age children can interpret the meanings of novel verbs as informed by the syntactic structures in a surprisingly detailed way.

This paradigm soon became a standard in studying Syntactic Bootstrapping in young children. Numerous experiments have adopted and extended this paradigm to explore different facets of Syntactic Bootstrapping. Many variations of this paradigm were developed: using pointing instead of looking as the behavioral responses measured (e.g., Arunachalam & Waxman, 2010; Kline, Snedeker, & Schulz, 2017; Rowland & Noble, 2010); using human actors as the protagonists of the actions instead of the humans in animal suits (e.g., Gertner & Fisher, 2012; Bunger & Lidz, 2006; Messenger, Yuan, & Fisher, 2015); adding a practice phase or a character-identification phase to reduce the task demands for young children (Gertner & Fisher, 2012; Scott, Chu, & Schulz, 2017; Scott & Fisher, 2009).

The changes can potentially bring unintended effects on the learning outcomes: studies differ significantly in the amount of training experience provided to the participants. Some studies, similar to the original Naigles (1990) design, distributed the training experience across multiple different novel verbs (e.g., He et al., 2020a; Arunachalam, 2013; Naigles & Kako, 1993; Naigles, 1996). Others, by the contrary, focused on training and testing on one novel verb throughout the experiment (e.g., Jin, 2015; Kline et al., 2017; Messenger et al., 2015). As a result, children may have heard the novel verb repeated from as little as 3 times (e.g. Naigles, 1990, 1996; Naigles & Kako, 1993), to as many as 27 times (e.g., Arunachalam, 2013) before they were tested. There are also changes motivated by theoretical considerations. For instance, whether the syntactic cues alone were sufficient to support verbs learning has been studied by altering the visual stimuli during the training phase. Rather than seeing the visual stimuli representing the relevant actions, the children in some training phases would see two actors talking to each other or one actor talking over the phone (e.g. Yuan & Fisher, 2009). The absence of extralinguistic cues did not prevent the children from interpreting the novel verbs correctly, which further suggests the significance of the syntactic cues in verb learning at an early age.

The changes in paradigms have also led to some inconsistent findings. For example, The results have been mixed on whether the children have a transitivity bias, that is, children are more likely to match the correct scenes when the novel verbs are embedded in the transitive frames. Some experiments found that children who were exposed to the transitive sentences would show a larger preference to the matching scenes, whereas children exposed to the intransitive sentences looked by chance or showed a looking pattern similar to the control group who were not provided with any relevant syntactic cues (Arunachalam & Waxman, 2010; Yuan et al., 2012). This transitivity advantage can be attributed to three factors: the clarity, the experience, and the perceptual saliency of the stimuli. First, under certain contexts, the transitive sentences have less ambiguity than their intransitive counterparts. If one sees two scenes: in one a girl patting a boy, in the other a girl and a boy jumping side-by-side, then upon hearing “the girl is gorping the boy”, only the scene with patting action is a plausible interpretation. In contrast, if one hears the intransitive sentence, “the girl is gorping”, the verb can be interpreted as being consistent with both the causative scene and the non-causative scene. Second, children may have more experience with transitive sentences than intransitive sentences. A corpus analysis on parental utterances for 1- to 6-year-old children found that transitive sentences make up approximately 24.36% of all utterances, whereas intransitive sentences occupy only 17.24% (Laakso & Smith, 2007). Thus the familiarity with transitive sentences can potentially improve children’s understanding of transitive sentences in the lab. Last but not least, the perceptual saliency of the visual scenes may have contributed to the asymmetrical performance in the transitive conditions and the transitive conditions. Children were found to have a baseline preference for two actors event over one actor event (e.g., Yuan et al., 2012), and for synchronous movement over causative movement (e.g., Naigles & Kako, 1993). These preferences can make detecting an effect in the transitive condition easier than in the intransitive condition.

However, other studies have found the opposite pattern. In the second and the third experiment in Naigles & Kako (1993), it was children in the transitive conditions who showed no preference for neither of the events, whereas children in the intransitive conditions looked longer at the matching event. Similarly, Bunger & Lidz (2004) also showed that children exposed to the intransitive sentences tend to look longer at the matching scene but those exposed to the transitive sentences look by chance. There were two proposed explanations for this reversed effect. First, transitive verbs can be more ambiguous under certain contexts due to children’s sensitivity to the internal structure of the event. Research has shown that children under 1 year of age already possess impressive abilities to detect the internal structure of an event. Following the structure, they can parse the event into subcomponents(Hespos, Grossman, & Saylor, 2010; Hespos, Saylor, & Grossman, 2009; Stahl, Romberg, Roseberry, Golinkoff, & Hirsh-Pasek, 2014; for a recent review, see Levine, Buchsbaum, Hirsh-Pasek, & Golinkoff, 2019). On the one hand, successful parsing underlies successful verb learning (Friend & Pace, 2011). But on the other hand, this sensitivity to the internal structure of the event can also lead to challenges for learning transitive verbs. As Bunger & Lidz (2004) pointed out, the transitive verbs can denote both the means of the action and the results of the action. In comparison, the possible meanings of intransitive verbs are more constrained to the results of the action. The wider range of possible meanings could make it more difficult for the children to correctly match the sentences and the scenes. The second possibility for this asymmetry is the processing demands brought by the words in the argument positions. Lidz et al. (2009) noted that the minimal demand from a simple transitive sentence is for children to map at least two nouns, the subject and the object to the corresponding agent and the patient. But for an intransitive sentence, the demand can be as low as only mapping only the subject of the sentence to the agent of the scene. In consequence, learning the novel verbs in the intransitive sentences can be easier than in the transitive sentences.

Besides the number of words in the verbs' arguments, the semantic content is also considered to be a key moderator for the learning process. Some scholars have proposed that semantically rich contexts are beneficial for verb learning, because the semantic content of the surrounding nouns can scaffold the interpretation of the verbs, together with the syntactic structure (Arunachalam & Waxman, 2011, 2015; Fisher, Hall, Rakowitz, & Gleitman, 1994; Gillette, Gleitman, Gleitman, & Lederer, 1999; Gleitman et al., 2005; Imai et al., 2008; Piccin & Waxman, 2007). But the empirical support for this view is mixed. In a recent study, He et al. (2020b) manipulated the amount of semantic content in the subject argument. Some preschoolers were provided with more information about the subject ("The tall girl is fezzing"), and others less ("The girl is fezzing"). Contrary to the semantic-scaffolding view, He et al. (2020b) found that the children learned better when the nouns were not modified, indicating that the children's limited information processing abilities may have impeded them from utilizing the semantic content. This pattern also persists when the number of words in the agent argument is controlled for. Pronouns have less semantic content than concrete nouns. Multiple studies have found that preschool children learn verbs better when the verbs co-occur with pronouns instead of nouns. They are not only more likely to correctly identify the scenes corresponding to the linguistic stimuli but also are more likely to generalize the syntactic frames to new verbs (Childers & Tomasello, 2001; Lidz et al., 2009).

To further shed light on these conflicting findings, we decided to conduct a meta-analysis on syntactic bootstrapping literature. Our first goal is to estimate the robustness of this phenomenon by calculating the meta-analytic effect size. Knowing the effect size is consequential to theory building (Bergmann et al., 2018; Lewis, 2016). Many past meta-analyses have tapped into other domains of early language development such as phonotactic learning (Cristia, 2018), word segmentation (Bergmann & Cristia, 2016), and mutual exclusivity (Lewis et al., 2020). To our knowledge, this is the first meta-analysis examining the development of syntactic abilities in early childhood and its influence on word learning. Our second goal is to examine the potential moderators of the effect. Understanding the moderators can be valuable for both theoretical reasons and practical reasons. Due to the limited sample sizes, individual infant experiments often have insufficient statistical power, which can make detecting relevant moderators difficult (Oakes, 2017). The variations among different testing procedures also made it difficult to reconcile the conflicting findings, because the differences in the observed effect can be caused by a combination of methodological factors, rather than reflecting the differences in the measured constructs. The meta-analytic approach has the opportunity to reveal some of the underlying interactions between different factors, which can inform future researchers about their experimental designs. Last but not least, our fourth goal is to evaluate how the effect changes with age. Most of the syntactic-bootstrapping studies used a cross-sectional design. Despite the virtue of being rigorously controlled, the “time-slice” nature of this design made it difficult to reconstruct a continuous developmental trajectory.

The meta-analysis can provide unique insights into how children's abilities to incorporate syntactic information into verbs learning to develop in the first three years of life.

References

Abbot-Smith, K., Imai, M., Durrant, S., & Nurmsoo, E. (2017). The role of timing and prototypical causality on how preschoolers fast-map novel verb meanings. *First Language*, *37*(2), 186–204.

* Arunachalam, S. (2013). Two-year-olds can begin to acquire verb meanings in socially impoverished contexts. *Cognition*, *129*(3), 569–573.

* Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, *114*(3), 442–446.

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

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.

Au, T. K.-F., Dapretto, M., & Song, Y.-K. (1994). Input vs constraints: Early word acquisition in korean and english. *Journal of Memory and Language*, *33*(5), 567–582.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J., … Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, (21), 85–123.

Bergmann, C., & Cristia, A. (2016). Development of infants' segmentation of words from native speech: A meta-analytic approach. *Developmental Science*, *19*(6), 901–917.

Bergmann, C., Tsuji, S., Piccinini, P. E., Lewis, M. L., Braginsky, M., Frank, M. C., & Cristia, A. (2018). Promoting replicability in developmental research through meta-analyses: Insights from language acquisition research. *Child Development*, *89*(6), 1996–2009.

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S.-Y., Pascual, L., … Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, dutch, french, hebrew, italian, korean, and american english. *Child Development*, *75*(4), 1115–1139.

Brown, R. W. (1957). Linguistic determinism and the part of speech. *The Journal of Abnormal and Social Psychology*, *55*(1), 1.

* 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). Cascadilla Press Somerville, MA.

* Bunger, A., & Lidz, J. (2006). Constrained flexibility in the acquisition of causative verbs. In *Proceedings of the annual boston university conference on language development* (Vol. 30, pp. 60–71).

Childers, J. B., Heard, M. E., Ring, K., Pai, A., & Sallquist, J. (2012). Children use different cues to guide noun and verb extensions. *Language Learning and Development*, *8*(3), 233–254.

Childers, J. B., & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the english transitive construction. *Developmental Psychology*, *37*(6), 739.

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, *38*(6), 967.

Childers, J. B., & Tomasello, M. (2006). 12 are nouns easier to learn than verbs? Three experimental studies. *Action Meets Word: How Children Learn Verbs*, 311.

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in korean: A cross-linguistic study. *Journal of Child Language*, *22*(3), 497–529.

Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. *Mechanisms of Language Acquisition. Hillsdale, NJ: Erlbaum*.

Cristia, A. (2018). Can infants learn phonology in the lab? A meta-analytic answer. *Cognition*, *170*, 312–327.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *Bmj*, *315*(7109), 629–634.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., … Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, i–185.

Fernald, A., Pinto, J. P., Swingley, D., Weinbergy, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, *9*(3), 228–231.

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*(1-4), 333–375.

Fisher, C., Jin, K. S., & Scott, R. M. (2020). The developmental origins of syntactic bootstrapping. *Topics in Cognitive Science*, *12*(1), 48–77.

Fort, M., Lammertink, I., Peperkamp, S., Guevara-Rukoz, A., Fikkert, P., & Tsuji, S. (2018). Symbouki: A meta-analysis on the emergence of sound symbolism in early language acquisition. *Developmental Science*, *21*(5), e12659.

Frank, M., Braginsky, M., Marchman, V., & Yurovsky, D. (in press). Variability and consistency in early language learning: The wordbank project.

Frank, M. C., Lewis, M. L., & MacDonald, K. (2016). A performance model for early word learning. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*.

Friend, M., & Pace, A. (2011). Beyond event segmentation: Spatial-and social-cognitive processes in verb-to-action mapping. *Developmental Psychology*, *47*(3), 867.

Furukawa, T. A., Barbui, C., Cipriani, A., Brambilla, P., & Watanabe, N. (2006). Imputing missing standard deviations in meta-analyses can provide accurate results. *Journal of Clinical Epidemiology*, *59*(1), 7–10.

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

Gentner, D. (2006). Why verbs are hard to learn. *Action Meets Word: How Children Learn Verbs*, 544–564.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. *Language Acquisition and Conceptual Development*, *3*, 215–256.

* Gertner, Y., & Fisher, C. (2012). Predicted errors in children's early sentence comprehension. *Cognition*, *124*(1), 85–94.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*(2), 135–176.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*(1), 3–55.

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

Goldin-Meadow, S., Seligman, M. E., & Gelman, R. (1976). Language in the two-year old. *Cognition*, *4*(2), 189–202.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*(1), 23–45.

* He, A. X., Huang, S., Waxman, S., & Arunachalam, S. (2020a). Two-year-olds consolidate verb meanings during a nap. *Cognition*, *198*, 104205.

He, A. X., Kon, M., & Arunachalam, S. (2020b). Linguistic context in verb learning: Less is sometimes more. *Language Learning and Development*, *16*(1), 22–42.

Hespos, S. J., Grossman, S. R., & Saylor, M. M. (2010). Infants' ability to parse continuous actions: Further evidence. *Neural Networks*, *23*(8-9), 1026–1032.

Hespos, S. J., Saylor, M. M., & Grossman, S. R. (2009). Infants' ability to parse continuous actions. *Developmental Psychology*, *45*(2), 575.

* Hirsh-Pasek, K., Golinkoff, R. M., & Naigles, L. (1996). Young children's use of syntactic frames to derive meaning. *The Origins of Grammar: Evidence from Early Language Comprehension*, 123–158.

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.

* Jin, K. S. (2015). *The role of syntactic and discourse information in verb learning* (PhD thesis). University of Illinois at Urbana-Champaign.

Kim, M., McGregor, K. K., & Thompson, C. K. (2000). Early lexical development in english-and korean-speaking children: Language-general and language-specific patterns. *Journal of Child Language*, *27*(2), 225–254.

* Kline, M., Snedeker, J., & Schulz, L. (2017). Linking language and events: Spatiotemporal cues drive children's expectations about the meanings of novel transitive verbs. *Language Learning and Development*, *13*(1), 1–23.

Laakso, A., & Smith, L. B. (2007). Pronouns and verbs in adult speech to children: A corpus analysis. *Journal of Child Language*, *34*(4), 725.

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

Levine, D., Buchsbaum, D., Hirsh-Pasek, K., & Golinkoff, R. M. (2019). Finding events in a continuous world: A developmental account. *Developmental Psychobiology*, *61*(3), 376–389.

Lewis, B., M. (2016). A quantitative synthesis of early language acquisition using meta-analysis. arXiv.

Lewis, M., Cristiano, V., Lake, B. M., Kwan, T., & Frank, M. C. (2020). The role of developmental change and linguistic experience in the mutual exclusivity effect. *Cognition*, *198*, 104191.

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

Lieven, E. V., Pine, J. M., & Barnes, H. D. (1992). Individual differences in early vocabulary development: Redefining the referential-expressive distinction. *Journal of Child Language*, *19*(2), 287–310.

Longobardi, E., Spataro, P., Putnick, D. L., & Bornstein, M. H. (2017). Do early noun and verb production predict later verb and noun production? Theoretical implications. *Journal of Child Language*, *44*(2), 480.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*(1), 57–77.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*.

Mathur, M., & VanderWeele, T. (2020). Sensitivity analysis for publication bias in meta-analyses.

* Messenger, K., Yuan, S., & Fisher, C. (2015). Learning verb syntax via listening: New evidence from 22-month-olds. *Language Learning and Development*, *11*(4), 356–368.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *PLoS Med*, *6*(7), e1000097.

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

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

* Naigles, L. R. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition*, *58*(2), 221–251.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, 1–135.

Nelson, K., Hampson, J., & Shaw, L. K. (1993). Nouns in early lexicons: Evidence, explanations and implications. *Journal of Child Language*, *20*(1), 61–84.

Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy*, *22*(4), 436–469.

Oviatt, S. L. (1980). The emerging ability to comprehend language: An experimental approach. *Child Development*, 97–106.

Papaeliou, C. F., & Rescorla, L. (2011). Vocabulary development in greek children: A cross-linguistic comparison using the language development survey. *Journal of Child Language*, *38*(4), 861.

Piccin, T. B., & Waxman, S. R. (2007). Why nouns trump verbs in word learning: New evidence from children and adults in the human simulation paradigm. *Language Learning and Development*, *3*(4), 295–323.

Rabagliati, H., Ferguson, B., & Lew-Williams, C. (2019). The profile of abstract rule learning in infancy: Meta-analytic and experimental evidence. *Developmental Science*, *22*(1), e12704.

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(6), 906–914.

* Rowland, C. F., & Noble, C. L. (2010). The role of syntactic structure in children's sentence comprehension: Evidence from the dative. *Language Learning and Development*, *7*(1), 55–75.

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*(5489), 265–266.

Schwartz, R. G., & Leonard, L. B. (1984). Words, objects, and actions in early lexical acquisition. *Journal of Speech, Language, and Hearing Research*, *27*(1), 119–127.

* Scott, K., Chu, J., & Schulz, L. (2017). Lookit (part 2): Assessing the viability of online developmental research, results from three case studies. *Open Mind*, *1*(1), 15–29.

* 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.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568.

Stahl, A. E., Romberg, A. R., Roseberry, S., Golinkoff, R. M., & Hirsh-Pasek, K. (2014). Infants segment continuous events using transitional probabilities. *Child Development*, *85*(5), 1821–1826.

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from mandarin speakers' early vocabularies. *Developmental Psychology*, *32*(3), 492.

Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the "noun bias" in context: A comparison of english and mandarin. *Child Development*, *70*(3), 620–635.

Tomasello, M. (2010). *Origins of human communication*. MIT press.

Viechtbauer, W. (2010). Conducting meta-analyses in r with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414–420.

* Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby?" Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science*, *20*(5), 619–626.

* Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: Simple structural cues

to verb meaning. *Child Development*, *83*(4), 1382–1399.

Abbot-Smith, K., Imai, M., Durrant, S., & Nurmsoo, E. (2017). The role of timing and prototypical causality on how preschoolers fast-map novel verb meanings. *First Language*, *37*(2), 186–204.

* Arunachalam, S. (2013). Two-year-olds can begin to acquire verb meanings in socially impoverished contexts. *Cognition*, *129*(3), 569–573.

* Arunachalam, S., & Waxman, S. R. (2010). Meaning from syntax: Evidence from 2-year-olds. *Cognition*, *114*(3), 442–446.

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

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.

Au, T. K.-F., Dapretto, M., & Song, Y.-K. (1994). Input vs constraints: Early word acquisition in korean and english. *Journal of Memory and Language*, *33*(5), 567–582.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J., … Hartung, J. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language*, (21), 85–123.

Bergmann, C., & Cristia, A. (2016). Development of infants' segmentation of words from native speech: A meta-analytic approach. *Developmental Science*, *19*(6), 901–917.

Bergmann, C., Tsuji, S., Piccinini, P. E., Lewis, M. L., Braginsky, M., Frank, M. C., & Cristia, A. (2018). Promoting replicability in developmental research through meta-analyses: Insights from language acquisition research. *Child Development*, *89*(6), 1996–2009.

Bornstein, M. H., Cote, L. R., Maital, S., Painter, K., Park, S.-Y., Pascual, L., … Vyt, A. (2004). Cross-linguistic analysis of vocabulary in young children: Spanish, dutch, french, hebrew, italian, korean, and american english. *Child Development*, *75*(4), 1115–1139.

Brown, R. W. (1957). Linguistic determinism and the part of speech. *The Journal of Abnormal and Social Psychology*, *55*(1), 1.

* 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). Cascadilla Press Somerville, MA.

* Bunger, A., & Lidz, J. (2006). Constrained flexibility in the acquisition of causative verbs. In *Proceedings of the annual boston university conference on language development* (Vol. 30, pp. 60–71).

Childers, J. B., Heard, M. E., Ring, K., Pai, A., & Sallquist, J. (2012). Children use different cues to guide noun and verb extensions. *Language Learning and Development*, *8*(3), 233–254.

Childers, J. B., & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the english transitive construction. *Developmental Psychology*, *37*(6), 739.

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology*, *38*(6), 967.

Childers, J. B., & Tomasello, M. (2006). 12 are nouns easier to learn than verbs? Three experimental studies. *Action Meets Word: How Children Learn Verbs*, 311.

Choi, S., & Gopnik, A. (1995). Early acquisition of verbs in korean: A cross-linguistic study. *Journal of Child Language*, *22*(3), 497–529.

Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. *Mechanisms of Language Acquisition. Hillsdale, NJ: Erlbaum*.

Cristia, A. (2018). Can infants learn phonology in the lab? A meta-analytic answer. *Cognition*, *170*, 312–327.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *Bmj*, *315*(7109), 629–634.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., … Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, i–185.

Fernald, A., Pinto, J. P., Swingley, D., Weinbergy, A., & McRoberts, G. W. (1998). Rapid gains in speed of verbal processing by infants in the 2nd year. *Psychological Science*, *9*(3), 228–231.

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*(1-4), 333–375.

Fisher, C., Jin, K. S., & Scott, R. M. (2020). The developmental origins of syntactic bootstrapping. *Topics in Cognitive Science*, *12*(1), 48–77.

Fort, M., Lammertink, I., Peperkamp, S., Guevara-Rukoz, A., Fikkert, P., & Tsuji, S. (2018). Symbouki: A meta-analysis on the emergence of sound symbolism in early language acquisition. *Developmental Science*, *21*(5), e12659.

Frank, M., Braginsky, M., Marchman, V., & Yurovsky, D. (in press). Variability and consistency in early language learning: The wordbank project.

Frank, M. C., Lewis, M. L., & MacDonald, K. (2016). A performance model for early word learning. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society*.

Friend, M., & Pace, A. (2011). Beyond event segmentation: Spatial-and social-cognitive processes in verb-to-action mapping. *Developmental Psychology*, *47*(3), 867.

Furukawa, T. A., Barbui, C., Cipriani, A., Brambilla, P., & Watanabe, N. (2006). Imputing missing standard deviations in meta-analyses can provide accurate results. *Journal of Clinical Epidemiology*, *59*(1), 7–10.

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

Gentner, D. (2006). Why verbs are hard to learn. *Action Meets Word: How Children Learn Verbs*, 544–564.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. *Language Acquisition and Conceptual Development*, *3*, 215–256.

* Gertner, Y., & Fisher, C. (2012). Predicted errors in children's early sentence comprehension. *Cognition*, *124*(1), 85–94.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*(2), 135–176.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*(1), 3–55.

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

Goldin-Meadow, S., Seligman, M. E., & Gelman, R. (1976). Language in the two-year old. *Cognition*, *4*(2), 189–202.

Golinkoff, R. M., Hirsh-Pasek, K., Cauley, K. M., & Gordon, L. (1987). The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*(1), 23–45.

* He, A. X., Huang, S., Waxman, S., & Arunachalam, S. (2020a). Two-year-olds consolidate verb meanings during a nap. *Cognition*, *198*, 104205.

He, A. X., Kon, M., & Arunachalam, S. (2020b). Linguistic context in verb learning: Less is sometimes more. *Language Learning and Development*, *16*(1), 22–42.

Hespos, S. J., Grossman, S. R., & Saylor, M. M. (2010). Infants' ability to parse continuous actions: Further evidence. *Neural Networks*, *23*(8-9), 1026–1032.

Hespos, S. J., Saylor, M. M., & Grossman, S. R. (2009). Infants' ability to parse continuous actions. *Developmental Psychology*, *45*(2), 575.

* Hirsh-Pasek, K., Golinkoff, R. M., & Naigles, L. (1996). Young children's use of syntactic frames to derive meaning. *The Origins of Grammar: Evidence from Early Language Comprehension*, 123–158.

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.

* Jin, K. S. (2015). *The role of syntactic and discourse information in verb learning* (PhD thesis). University of Illinois at Urbana-Champaign.

Kim, M., McGregor, K. K., & Thompson, C. K. (2000). Early lexical development in english-and korean-speaking children: Language-general and language-specific patterns. *Journal of Child Language*, *27*(2), 225–254.

* Kline, M., Snedeker, J., & Schulz, L. (2017). Linking language and events: Spatiotemporal cues drive children's expectations about the meanings of novel transitive verbs. *Language Learning and Development*, *13*(1), 1–23.

Laakso, A., & Smith, L. B. (2007). Pronouns and verbs in adult speech to children: A corpus analysis. *Journal of Child Language*, *34*(4), 725.

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

Levine, D., Buchsbaum, D., Hirsh-Pasek, K., & Golinkoff, R. M. (2019). Finding events in a continuous world: A developmental account. *Developmental Psychobiology*, *61*(3), 376–389.

Lewis, B., M. (2016). A quantitative synthesis of early language acquisition using meta-analysis. arXiv.

Lewis, M., Cristiano, V., Lake, B. M., Kwan, T., & Frank, M. C. (2020). The role of developmental change and linguistic experience in the mutual exclusivity effect. *Cognition*, *198*, 104191.

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

Lieven, E. V., Pine, J. M., & Barnes, H. D. (1992). Individual differences in early vocabulary development: Redefining the referential-expressive distinction. *Journal of Child Language*, *19*(2), 287–310.

Longobardi, E., Spataro, P., Putnick, D. L., & Bornstein, M. H. (2017). Do early noun and verb production predict later verb and noun production? Theoretical implications. *Journal of Child Language*, *44*(2), 480.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*(1), 57–77.

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*.

Mathur, M., & VanderWeele, T. (2020). Sensitivity analysis for publication bias in meta-analyses.

* Messenger, K., Yuan, S., & Fisher, C. (2015). Learning verb syntax via listening: New evidence from 22-month-olds. *Language Learning and Development*, *11*(4), 356–368.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The prisma statement. *PLoS Med*, *6*(7), e1000097.

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

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

* Naigles, L. R. (1996). The use of multiple frames in verb learning via syntactic bootstrapping. *Cognition*, *58*(2), 221–251.

Nelson, K. (1973). Structure and strategy in learning to talk. *Monographs of the Society for Research in Child Development*, 1–135.

Nelson, K., Hampson, J., & Shaw, L. K. (1993). Nouns in early lexicons: Evidence, explanations and implications. *Journal of Child Language*, *20*(1), 61–84.

Oakes, L. M. (2017). Sample size, statistical power, and false conclusions in infant looking-time research. *Infancy*, *22*(4), 436–469.

Oviatt, S. L. (1980). The emerging ability to comprehend language: An experimental approach. *Child Development*, 97–106.

Papaeliou, C. F., & Rescorla, L. (2011). Vocabulary development in greek children: A cross-linguistic comparison using the language development survey. *Journal of Child Language*, *38*(4), 861.

Piccin, T. B., & Waxman, S. R. (2007). Why nouns trump verbs in word learning: New evidence from children and adults in the human simulation paradigm. *Language Learning and Development*, *3*(4), 295–323.

Rabagliati, H., Ferguson, B., & Lew-Williams, C. (2019). The profile of abstract rule learning in infancy: Meta-analytic and experimental evidence. *Developmental Science*, *22*(1), e12704.

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(6), 906–914.

* Rowland, C. F., & Noble, C. L. (2010). The role of syntactic structure in children's sentence comprehension: Evidence from the dative. *Language Learning and Development*, *7*(1), 55–75.

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*(5489), 265–266.

Schwartz, R. G., & Leonard, L. B. (1984). Words, objects, and actions in early lexical acquisition. *Journal of Speech, Language, and Hearing Research*, *27*(1), 119–127.

* Scott, K., Chu, J., & Schulz, L. (2017). Lookit (part 2): Assessing the viability of online developmental research, results from three case studies. *Open Mind*, *1*(1), 15–29.

* 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.

Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568.

Stahl, A. E., Romberg, A. R., Roseberry, S., Golinkoff, R. M., & Hirsh-Pasek, K. (2014). Infants segment continuous events using transitional probabilities. *Child Development*, *85*(5), 1821–1826.

Tardif, T. (1996). Nouns are not always learned before verbs: Evidence from mandarin speakers' early vocabularies. *Developmental Psychology*, *32*(3), 492.

Tardif, T., Gelman, S. A., & Xu, F. (1999). Putting the "noun bias" in context: A comparison of english and mandarin. *Child Development*, *70*(3), 620–635.

Tomasello, M. (2010). *Origins of human communication*. MIT press.

Viechtbauer, W. (2010). Conducting meta-analyses in r with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48.

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414–420.

* Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby?" Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science*, *20*(5), 619–626.

* Yuan, S., Fisher, C., & Snedeker, J. (2012). Counting the nouns: Simple structural cues

to verb meaning. *Child Development*, *83*(4), 1382–1399.