A meta-analysis of the syntactic bootsrapping phenomenon in word learning

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:

Abstract

so abstract!

Keywords: keyword1

Word count:

A meta-analysis of the syntactic bootsrapping phenomenon in word learning

Methods

Literature Search

We conducted our literature search following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses checklist (PRISMA; Moher, Liberati, Tetzlaff, Altman, & Prisma Group, 2009). We identified relevant papers by conducting 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: $N = 3339$; retrieved between May 2020 and July 2020). We screened the first 60 pages (N = 600) of the keyword search results and the first 10 pages (N = 100) of the forward search results. 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, Jin, & Scott, 2020) and the experts in the field (N = 11). In our final sample, we included published journal articles, conference proceedings, doctoral dissertations, and unpublished manuscripts. They will be collectively referred to as "papers" in the following sections. Each paper may include multiple experimental conditions, and thus provides multiple effect sizes for the final analysis.

We restricted our final sample to papers that satisfied the following conditions: First, the experimental paradigm involved a two-alternative forced-choice task, in which the participants were instructed to identify the scene that matched the linguistic stimuli. Second, the visual stimuli were displayed on a monitor. We included videos that were either recordings of 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", while "Look, Kradding!" does not and therefore would be excluded. Finally, we resricted our sample to English-speaking, typically-developing children.

Our final sample included data from unique infants (Mean age: Days), reported in 106 individual effect sizes from 30 individual papers.

Data Entry

For each paper, we entered the paper metadata (e.g., citation), information to calculate effect sizes, and moderators. We entered a separate effect size for each experimental condition and age group in paper. Most papers therefore contained multiple effect sizes.

In order to calculate the effect sizes, we recorded the sample size of each condition, the group mean, and the across-participant standard deviation. The mean and standard deviation were obtained from one of the three ways: a) directly retrieved from the results section or the data-presenting tables (N = 71); b) recovered from the plots by measuring the height of the bars and the error bars (N =27); c) contacted the original authors (N = 8). For looking time studies, when the paper only reported the raw looking time in seconds, we calculated the proportion of correct response by dividing the mean looking time toward the matching scenes by the sum of mean looking time toward the matching scenes and the mean looking time toward the non-matching scenes (i.e., excluding the look away time from the denominator). The standard deviations were also scaled by being divided by the sum.

We classified our final samples of effect sizes (N = 106) into two categories: the prototypical ones (N =) and atypical ones (N =). Prototypical studies were the ones in which the set-up and the procedures were most similar to the seminal paper (Naigles, 1990). The atypical studies were experiments that met all of the inclusion criteria, but deviated from the prototypical designs in some non-trivial ways (The stimuli involved more complex syntactic structure or event structure: N = 94; The testing procedure included a contrast phase: N = 10; The experiment was conducted through online platform: N = 2). To preserve the homogeneity of the experiments without reducing the power of our analysis, we conducted our statistical analysis both on the entire sample and specifically on the prototypical studies.

For all experimental conditions, we coded three types of moderators: participant moderators, stimuli moderators, and procedure moderators. The information was either retrieved from the methods section of the paper or obtained via contacting authors.

At the participant level, we coded infants' mean age in days (N= 106) and the median productive vocabulary measured by MacArthur-Bates Communicative Development Inventories (CDI) Words and Sentences ().

We coded features of the linguistic stimuli during both training and testing. The training linguistic stimuli were categorized as transitive (N =), if the novel verbs were embedded in a sentence with two or more noun arguments, and intransitive (N =) if the novel verbs were embedded in a sentence with one noun argument. We also coded the types of words used in the agent argument (One noun: N = 31; One pronoun: N = 11; Two nouns: N = ; Noun phrase: N = 11; Varying across sentences: N = 26) and the patient argument(One noun: N = 35; One pronoun: N = 13; Noun phrase: N = 15; Varying across sentences: N = 4). We coded the testing linguistic stimuli for whether infants were prompted to identify the action (e.g. "Where's lorping? Find lorping!", N =) or the actors (e.g. "Which one (verbed) the other one….point!", N =).

For visual stimuli, we coded the types of media used (video or animation) and the types of protagonists in the events (person or non-person). 82 conditions used video recordings of human actors (N =) or human actors in animal suits (N =). The other 24 conditions used clips of animation with non-person figures as the protagonists of the events. We also coded how the onset of the linguistic stimuli aligned with the visual stimuli. The procedure was coded as "simultaneous" if the very first training sentence was presented along with the visual stimuli depicting relevant action (N =). It was coded as "immediately after" if it was presented along with an attention-getter or a blank screen, immediately followed by the relevant action (N =). Finally, some experimental conditions first presented the linguistic stimuli paired with irrelevant visual scenes (e.g. a person on the phone talking). The relevant visual stimuli were not shown until the training phase is over. For experimental conditions using this alignment, they were coded as "asynchronous"(N =).

Finally, procedure moderators included the types of response elicited from the participants: whether the infants were explicitly prompted to point or their eye gaze duration were measured as they heard the linguistic stimuli (Pointing: N = ; Looking: N =). Three characteristics of the experimental procedures were coded as categorical variables: the inclusion of practice phase (Yes: N = ; No: N =), the inclusion of character-identification phase (Yes: N = ; No: N =), and the distribution of the training and the testing trials (Mass: N = ; Distributed: N =). A procedure was categorized as "mass" if and only if the infants were trained exclusively on one novel verb and tested on the very same verb. It was "distributed" if the infants were given multiple train and test pairs on multiple novel verbs. To better characterize the experience infants had prior to testing, we also coded how many train-test pair the infants were given (for mass procedure it was always 1), how many trials during the test phase infants were given, how many times the visual stimuli showing the relevant actions were presented, and how many times each novel verb was spoken in a syntactically-informative way.

Data Preprocessing

We calculated the Cohen's d effect size for each experimental condition. Because we only included studies that used a two alternative forced-choice test method, we compared the mean proportion of correct response against chance level (0.5) for the group. Although some papers collected a baseline measure (N = ; e.g. XXX, XXX), most papers did not. In order to calculate a consistent effect size measure across all condition, we used chance level as a baseline for calculating effect size for all conditions (see SI for comparision of effect sizes using reported baseline versus chance).

All the effect sizes and the coded variables were then analyzed with the metafor package in R (Viechtbauer, 2010).

Results

A Section Name

General Discussion

References

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

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.

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

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

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.

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

Appendix