Title

The Mystery of Early Taxonomic Development

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN

1069-7977

Authors

Yim, Hyungwook
Savic, Olivera
Barkhimer, Alexandria
et al.

Publication Date

2021

Peer reviewed

The Mystery of Early Taxonomic Development

Hyungwook Yim (hwyim@hanyang.ac.kr)
Department of Cognitive Sciences, Hanyang University, Republic of Korea
School of Psychological Sciences, The University of Melbourne, Australia

Olivera Savic (savic.7@osu.edu)
Department of Psychology, The Ohio State University, USA

Alex Barkhimer (barkhimer.1@osu.edu)
Department of Psychology, The Ohio State University, USA

Vladimir M. Sloutsky (sloutsky.1@osu.edu)
Department of Psychology, The Ohio State University, USA

Simon J. Dennis (simon.dennis@unimelb.edu.au)
School of Psychological Sciences, The University of Melbourne, Australia

Abstract

Research has long investigated how knowledge about the world is connected by meaningful, semantic links. Much of this research has focused on a specific type of semantic link known as a taxonomic link, which connects concepts belonging to the same semantic category. However, many inconsistencies have emerged regarding *how* and *when* this specific type of semantic link is formed. The goal of the present study was to investigate this contradicting research and provide an explanation for the inconsistent results. To do this, we examined the linguistic environment of stimuli sets from three studies that either supported or did not support protracted taxonomic development. Results provided evidence that the idea that semantic links between members of taxonomic categories in early development may be based on simple co-occurrences in language.

Keywords: semantic development; semantic organization; taxonomic relation; computer simulation; corpus analysis

Introduction

Decades of research into semantic development has focused on understanding the emergence of one specific aspect of semantic organization: taxonomic links. Taxonomic links connect concepts that belong to the same stable semantic category, such as bird, fruit or furniture. The focus on taxonomic links is not surprising having in mind that they support one of the central facets of human intelligent behavior – our ability to extend knowledge from known to novel (Gelman & Markman, 1986; Inhelder & Piaget, 1964; Markman & Hutchinson, 1984; Osherson, Smith, Wilkie, Lopez, & Shafir, 1990). For example, upon learning that pigeons have hollow bones, we can infer that robins, storks and other birds are likely to have hollow bones as well. The two critical questions driving the research on taxonomic development are the questions of how and when human learners start to form taxonomic links. In what follows, we first review the current views on how taxonomic links could be formed. We then highlight a controversy in the field regarding when sensitivity to taxonomic links emerges: While some argue that taxonomic links may start forming before the age of two (e.g. Arias-Trejo & Plunkett, 2013; Bergelson, & Aslin, 2017), others suggest that taxonomic development is protracted till preschool age (e.g. Unger, Savic, & Sloutsky, 2020; Unger, Vales, & Fisher, 2020). Finally, we present and explicitly test an explanation that may provide a reconciliation for the two opposing views.

Emergence of taxonomic links: *How* and *when*

How do human learners form taxonomic links? A widely adopted and substantially supported perspective is that learners capitalize on the fact that members of taxonomic categories often reliably share observable features (Fisher, 2015; McClelland & Rogers, 2003; Quinn & Eimas, 1997; Rosch, 1978; Sloutsky, 2010). For example, birds have beaks, feathers and wings, while fish have gills, scales and fins. According to this perspective, grouping entities based on shared features forms a foundation for taxonomic organization.

Alternatively, taxonomic links can be formed based on the regularities in language (e.g. Sloutsky, Yim, Yao, & Dennis, 2017; Savic, Unger, & Sloutsky, 2020). Specifically, words denoting taxonomically related concepts reliably occur in similar language contexts. For example, *pigeon* and *stork* both reliably co-occur with words such as *fly* and *nest* in language, whereas *horse* and *donkey* both reliably co-occur with *ride* and *saddle* in language.

Although there are important differences in the two proposed learning mechanisms supporting taxonomic knowledge, both of these routes rely on cognitive structures that show protracted development. Specifically, selective attention mechanisms, which support formation of taxonomic links from shared features, only emerge around the age of six (Sloutsky, 2010; Deng & Sloutsky, 2016). Similarly, mechanisms needed for integration of information across overlapping linguistic contexts are immature even in four-year-old children (Sloutsky, Yim, Yao, & Dennis, 2017; Unger, Savic, & Sloutsky, 2020), and may require further maturation (Yim, Savic, Unger, Sloutsky, & Dennis, 2019). Therefore, taxonomic links formed based on the observance of shared features or shared language contexts are not expected to be formed before preschool age.

Recently, a number of studies have challenged the view of protracted taxonomic development, offering behavioral (e.g. Arias-Trejo & Plunkett, 2013; Bergelson, & Aslin, 2017) and electrophysiological evidence (Rämä, Sirri, & Serres, 2013) for taxonomic priming effects before the age of 24 months. In an influential study using an intermodal preferential looking task, 24-month-olds looked more to a target image (e.g. *bike*) when it was preceded by a taxonomically related word (e.g. *lorry*) than an unrelated word (e.g. *apple*) (Arias-Trejo & Plunkett, 2013). Supporting results were obtained in an experiment using a visual word paradigm (Chow, Davies, & Plunkett, 2017), where 24- and 30-month-olds, after hearing a target word (e.g. *bee*), preferentially fixated to the taxonomic distractor (e.g. *cat*) over the unrelated distractor (e.g. *dress*). Similarly, in a study examining event-related potentials, researchers found neural evidence of taxonomic semantic priming in 18- and 24-month-olds (Rämä, Sirri, & Serres, 2013). The most striking finding from this group is the evidence of semantic competition in 6-month-old infants raised by presence of taxonomically related items, such as *stroller-car, milk-juice or foot-hand* (Bergelson, & Aslin, 2017). Taken together, these findings suggest that even early in childhood, hearing a word activates taxonomically related words.

Potential explanation of early taxonomic priming

If 6- and 24-month-olds can not form links based on the overlap in shared features or shared linguistic contexts, how are early taxonomic links formed? One possibility is that early in development links between the members of the same taxonomic category may be formed based on a simple co-occurrence regularity. There are several arguments that make this assumption plausible.

First, although they are often contrasted, co-occurrence and taxonomic relations are orthogonal constructs. Many taxonomically related items also frequently co-occur in the environment. For example, farm animals such as *cow*, *horse*, *pig* or breakfast food such as *eggs*, *bacon*, *cereal* are taxonomically related but also co-occur frequently. Second, extensive computational literature suggests that everyday language input, including the input to young children, is abundant in co-occurrence regularities from which semantic links can be formed (e.g., Landauer, & Dumais, 1997). Finally, numerous statistical learning studies have demonstrated that even infants are sensitive to co-occurrence regularities in visual and auditory domains (Saffran, & Kirkham, 2018). In addition, a number of studies demonstrated that semantic relations based on co-occurrence (i.e. associative, thematic, schema) emerge early in development (for a recent review, see Unger, Vales, & Fisher, 2020).

Therefore, it is reasonable to expect that young learners can form links between some members of taxonomic categories. Specifically, they are expected to be able to form links between taxonomically related items that reliably and directly co-occur either in the visual environment or in language. On the other hand, taxonomic links that require observance of shared features or shared linguistic contexts are not expected to emerge before the preschool age.

The current study

Whereas some studies suggest that taxonomic relations emerge later in childhood (around the age of six years), other studies present evidence that even infants have formed taxonomic relations. The aim of the current study is to provide an explicit test of an explanation that may reconcile the inconsistent results.

We propose that the disagreement between these studies rise from the difference in the stimuli sets that were used. Although all studies examined taxonomically related word pairs, it is possible that the studies demonstrating the early emergence of taxonomic competence used word pairs that consisted of words that also frequently co-occurred in the linguistic environment. If this is true, such co-occurrence relations can explain the observed priming effects even in the absence of a taxonomic structure. The idea is worth examining because, if confirmed, it would eliminate an observed inconsistency between the studies finding evidence for the early emergence of taxonomic competence and those finding evidence for the late emergence.

We tested this idea by examining the linguistic environment (i.e., TASA corpus) of the word pairs used in all the studies. Then, we calculated how often the pairs co-occur and simulated the task of interest using a co-occurrence measure (i.e., cubed pointwise mutual information). We expect that the studies that provided evidence for the early taxonomic competence (i.e., taxonomic priming effect) will show a high measure of co-occurrence between the taxonomically related words within each pair. On the other hand, we expect that the studies that provide evidence for the late emergence will show a low measure of co-occurrence, similar to that between unrelated words.

Simulation

As different studies used different tasks, we tried to select behavioral studies that used a relatively similar task. Moreover, as we are not suggesting a process model, but examining the effects of association strength that stems from co-occurrence statistics in the linguistic environment, we also tried to select tasks that can assume a simple mechanism. The following studies have been selected.

Selected studies

Arias-Trejo & Plunkett (2013) The study supports a taxonomic priming effect. The task involves an auditory prime word followed by a target word. Then, an image of the target word and a distractor is presented side by side. Preferential looking time is used to evaluate the priming effect. There were two conditions, where in one condition the prime-target has a taxonomic relation and another condition the prime-target has an associative relation. The results support a priming effect in both conditions.

Bergelson, & Aslin (2017) The study supports a taxonomic priming effect. The task involved an auditory presentation of a target word followed by an image of the target word and a distractor. The distractor can be either unrelated or taxonomically related to the target word (i.e., unrelated trials or taxonomic trials). The results show that in the unrelated trials, looking time at the target image was longer than chance. On the other hand, in the taxonomic trials, looking time at the target image was not longer than chance. The results are interpreted as evidence for the existence of a taxonomic relation as the taxonomically related words (in the taxonomic trials) distracted the infants from looking at the target image.

Unger, Savic, & Sloutsky (2020) The study does not find evidence for a taxonomic priming effect. There are three studies that provide the same conclusion. However, we will focus on Experiment 2 as the task is similar to the tasks in the other studies. The task involved an auditory presentation of the target word followed by an image that (1) matched the target word, (2) associatively related to the target word, (3) taxonomically related to the target word, or (4) unrelated to the target word. The main analysis concerned how distracted the participants become in the non-matching trials (i.e., where a No response is correct; correct rejection) using reaction time measures. Results showed that it took longer to reject the associatively related words compared to the unrelated word, whereas there was no difference in rejecting the taxonomically related words compared to the unrelated words. The results are interpreted as there is no evidence for a taxonomic relation.

Co-occurrence measure

We calculated the cubed pointwise mutual information (PMI^3) using the TASA corpus (Touchstone Applied Science Associates, Inc.). PMI^3 has been shown to be robust in measuring co-occurrence-based word similarity compared to the traditional PMI measure, which tends to give higher values for low-frequency words (Role, & Nadif, 2011). PMI^3 is calculated as shown in Equation 1.

$$PMI^3(A,B) = log_2(\frac{p(A,B)^3}{p(A) \cdot p(B)}) \quad (1)$$

, where $p(A)$ and $p(B)$ are the probabilities of word A and word B occurring in the corpus respectively. $p(A,B)$ is the probability of word A and word B co-occurring. In particular, $p(A,B)$ was calculated as the probability of word B occurring within a certain window before and after the word A. We examined a window size of ±3, ±7, ±11, and ±15 words. The PMI^3 values were then converted into percentile scores based on the empirical cumulative distribution that was derived from all PMI^3 values calculated from the corpus. These percentile scores tell how strong the co-occurence strengths are between two words among all word pairs that co-occurred. Depending on the study, the percentile scores were processed differently as follows:

Arias-Trejo & Plunkett (2013) As the dependent measure of the task was the looking time between an image of the target word and a distractor, we compared the percentile score between the prime and the target and the prime and the distractor using the Luce choice rule (Luce, 1959) as in Equation 2.

$$P(Target) = \frac{w_t}{(w_t + w_d)} \quad (2)$$

, where $P(Target)$ is the proportion of looking to the target image, w_t is the percentile score between the prime and target, and w_d is the percentile score between the prime and distractor.

Bergelson, & Aslin (2017) The dependent measure of the task was between two different distractors – a taxonomically related one or an unrelated one. As these two distractors are not tested together we did not use a Luce choice rule. Instead we directly compared the percentile score between the target word and the taxonomic distractor and the percentile score between the target word and the unrelated distractor (note that unlike Arias-Trejo & Plunkett (2013) there was no prime word in this task).

Unger, Savic, & Sloutsky (2020) The task structure is similar to Bergelson, & Aslin (2017). However, there are three conditions – taxonomically related distractor, associatively related distractor, and unrelated distractor. Therefore, we calculated the percentile score between the target word and these three kinds of distractors respectively.

Results

Simulation results when using the words in Arias-Trejo & Plunkett (2013) are shown in Figure 1A. The taxonomic condition (Figure 1A left panel) showed above chance values when using a window size of ±7 (p = .025, two-tailed t-test against .5), ±11 (p = .002), and ±15 words (p = .003), but not when using a size of ±3 words (p = .40; note that the proportion of looking values were calculated using a Luce choice rule for this task as there was a distractor at test). The associative condition (Figure 1A right panel) also showed above chance values when using a window size of ±7 (p = .003), ±11 (p = .020), and ±15 words (p = .030), but not when using a size of ±3 words (p = .12). The unrelated trials for both conditions all did not show above chance values for all window sizes ($p > .10$).

Simulation results when using the words in Bergelson, & Aslin (2017) are shown in Figure 1B. The taxonomically related words showed higher percentile scores than the unrelated words when using a window size of ±7 ($p < .001$; two-tailed t-test between taxonomically related and unrelated

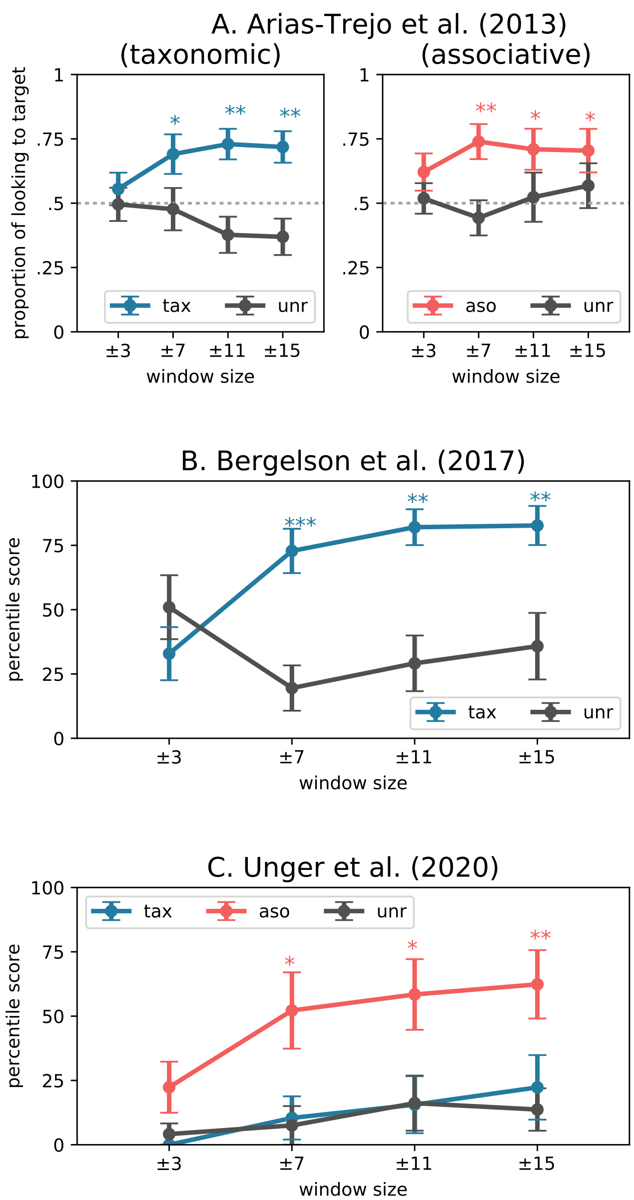

Figure 1: Simulation results. (A) results from Arias-Trejo et al. (2013) the left panel shows results from the taxonomic priming condition (tax) while the right panel shows results from the associative priming condition (aso) with its corresponding unrelated conditions (unr). The values show the simulated proportion of looking at the target in the task using the Luce choice rule, where .5 represents chance level, (B) results from Bergelson et al. (2017), where the taxonomic (tax) and unrelated (unr) distractor conditions are presented, and (C) results from Unger et al. (2020), where the taxonomic (tax), associative (aso), and unrelated (unr) distractor conditions are presented. Error bars represent ±1SEM. *, **, *** represents $p < .05, p < .01, p < .001$ respectively.

conditions), ±11 ($p = .001$), and ±15 words ($p = .009$), but not when using a size of ±3 words ($p = .30$).

Results generated using the words from Unger, Savic, & Sloutsky (2020) are shown in Figure 1C. We took the unrelated condition as the baseline and examined the difference between the taxonomic and unrelated condition, and between the associative condition and the unrelated condition. Only the associative condition showed a difference when using a window size of ±7 ($p = .020$), ±11 ($p = .028$), and ±15 words ($p = .008$), but not when using a size of ±3 words ($p = .12$). The taxonomic conditions did not show a difference across all window sizes ($p > .1$).

In sum, the results supported our hypothesis where the existence or non-existence of the priming effect is driven by the co-occurence measure (i.e., percentile values of the cubed pointwise mutual information). The studies that showed a priming effect such as Arias-Trejo & Plunkett (2013) and Bergelson, & Aslin (2017) showed a high measure of co-occurrence between the target and the taxonomic prime word (or the distractor that acted as a prime word). On the other hand, the study that did not find the effect (i.e., Unger, Savic, & Sloutsky, 2020) did not show a high measure of co-occurrence between the target and the taxonomic prime word.

Additionally, an interesting finding is that the results were only shown when using a window size wider than ±7 words, and not when the windows size was ±3 words. This possibly provides information about the properties of the co-occurrence similarity that is required to form semantic relations. We will further discuss this finding in the discussion section.

Discussion

A number of recent studies have reported taxonomic competence in children younger than two-years of age. These findings stand in stark contrast with our current understanding of the development of the learning mechanism supporting the formation of taxonomic links. The present study offers a potential explanation for this discrepancy.

Our simulations demonstrate that in studies reporting early taxonomic competence, taxonomically related items more reliably co-occur in the linguistic environment compared to the unrelated items. In contrast, in a study where related and unrelated items did not differ on the measure of simple co-occurrence, taxonomic priming was found to be weak even in four-year-olds. Therefore, the present study offers supporting evidence for the assumption that early in development, semantic links between the members of the same taxonomic category may be based on simple co-occurrence regularities in language. Our findings further suggest that some of the interpretations of the early taxonomic competence need to be reconsidered. Here we discuss what can be learned from these studies and offer suggestions for future work.

What kind of co-occurrence is required to form taxonomic relations?

The simulation results showed that some of the taxonomic relations can be formed through simple co-occurrence. Among the four window sizes that we examined, a window size of ±7, ±11, and ±15 words captured the relationship between the word pairs, whereas a window size of ±3 words did not. This result is worth noting as it may have important theoretical implications. The results suggest that the linguistic environment that is required for forming early taxonomic relations is not any co-occurrence. From the current study we can at least rule out that the nearby co-occurrence statistics (i.e., three words before and after the target word) does not explain the formation of early taxonomic relation. This window size of ±3 words may possibly include more grammatical relations and words that have different parts of speech, whereas taxonomically related words typically belong to the same grammatical class (e.g. nouns) and thus would not co-occur often in a narrow window. For example, in a sentence "the big black ***dog*** ran to the fence", the words that co-occurs with the target word in a ±3-words window are *the*, *big*, *black*, *ran*, *to*, and *the*, which are all not taxonomically related to 'dog' but provide grammatical information (e.g. the-dog) or thematic/associative information (e.g., big-dog, black-dog). Therefore, the information that is required to form taxonomic relations requires a wider window size, which is, based on the current study, wider than ±7-words. On the other hand, further examination is required on how wide a window size contributes to forming early taxonomic relations both regarding the actual information that contributes and regarding the developmental cognitive constraints.

How the analyses of language input may advance research on semantic development

Semantic development research typically involves measuring the degree to which concepts that adults judge to be semantically linked in a specified way (e.g. taxonomic) are linked in children. In addition to being contaminated by other kinds of semantic and non-semantic relations (e.g. Bassok & Medin, 1997), adult generated norms tell us little about how children may form semantic links. One possible solution to this problem, as demonstrated here, may be in the selection of stimuli based on the regularities in the environment. By offering a more objective measure of a link between the concepts, this approach could eliminate inconsistencies stemming from differences in stimuli selection. At the same time, measures of environmental regularities can advance our understanding of the mechanisms supporting formation of semantic links and thus go a step further than a mere description of developmental changes in semantic organisation.

Future directions

To the best of our knowledge, the current study represents the first attempt to explicitly test the role of direct co-occurrence in explaining the recently reported early taxonomic priming effects. While this study makes an important step forward, further work will be needed to test whether the explanation provided here generalises to a wider selection of studies that examine semantic development. In addition, as we mentioned above, our understanding of corpus measures is still developing. In addition to important technical aspects such as the window size and the choice of the corpus (e.g. TASA vs. CHILDES, MacWhinney, 2000), there are important measures of the environmental regularities which contribution to taxonomic development is worth considering. For example, in addition to the measures of direct co-occurrence, measures of shared context (i.e. the overlap in patterns of direct co-occurrence) may play a critical role in taxonomic development (Sloutsky, Yim, Yao, & Dennis, 2017; Unger, Savic, & Sloutsky, 2020).

Conclusion

The current study provides an explicit test of the assumption that early in development, links between the items of the same taxonomic category may be based on simple co-occurrence regularities rather than the overlap in features or overlap in language contexts. Our simulations provide support for this assumption: In studies reporting early taxonomic priming, taxonomically related items more reliably co-occur in the linguistic environment compared to the unrelated items. These findings suggest that some of the interpretations of the early taxonomic priming effects may need to be reconsidered.

Acknowledgments

This research was supported under Australian Research Council's (ARC) Discovery Projects funding scheme to SJD (DP150100272), and the National Institutes of Health Grants R01HD078545 and P01HD080679 to VMS.

References

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

Bassok, M., & Medin, D. (1997). Birds of a feather flock together: Similarity judgments with semantically rich stimuli. *Journal of Memory and Language*, *36*, 311–336.

Bergelson, E., & Aslin, R. (2017). Nature and origins of the lexicon in 6-mo-olds. *Proceedings of the National Academy of Sciences*, *114*(49), 12916–21.

Chow, J., Davies, A., & Plunkett, K. (2017). Spoken-word recognition in 2-year-olds: The tug of war between phonological and semantic activation. *Journal of Memory and Language*, *93*, 104–134.

Deng, W., & Sloutsky, V. (2016). Selective attention, diffused attention, and the development of categorization. *Cognitive Psychology*, *91*, 24–62.

Fisher, A. (2015). Development of inductive generalization. *Child Development Perspectives*, *9*, 172–177.

Gelman, S., & Markman, E. (1986). Categories and induction in young children. *Cognition*, *23*, 183–209.

Goldstone, R., & Rogosky, B. (2002). Using relations within conceptual systems to translate across conceptual systems. *Cognition*, *84*(3), 295–320.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child*. New York: Norton.

Landauer, T., & Dumais, S. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, *104*(2), 211–240.

Luce, R. (1959). *Individual choice behavior: A theoretical analysis*. New York: Wiley.

MacWhinney, B. (2000). *The childes project: The database* (Vol. 2). Psychology Press.

Markman, E., & Hutchinson, J. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, *16*, 1–27.

McClelland, J., & Rogers, T. (2003). The parallel distributed processing approach to semantic cognition. *Nature Reviews Neuroscience*, *4*, 310–322.

Osherson, D., Smith, E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, *97*, 185.

Quinn, P., & Eimas, P. (1997). A reexamination of the perceptual-to-conceptual shift in mental representations. *Review of General Psychology*, *1*, 271–287.

Role, F., & Nadif, M. (2011). Handling the impact of low frequency events on co-occurrence based measures of word similarity: A case study of pointwise mutual information. In January (Ed.), *Kdir 2011 - proceedings of the international conference on knowledge discovery and information retrieval* (p. 226–231).

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (p. 27–48). Hillsdale, NJ: Lawrence Erbaum Associates.

Rämä, P., Sirri, L., & Serres, J. (2013). Development of lexical–semantic language system: N400 priming effect for spoken words in 18-and 24-month old children. *Brain and Language*, *125*(1), 1–10.

Saffran, J., & Kirkham, N. (2018). Infant statistical learning. *Annual Review of Psychology*, *69*(1), 181–203.

Savic, O., Unger, L., & Sloutsky, V. (2020). Becoming organized: How simple learning mechanisms may shape the development of rich semantic knowledge. In *Proceedings of the 42nd annual conference of the cognitive science society.*

Sloutsky, V. (2010). From perceptual categories to concepts: What develops? *Cognitive Science*, *34*, 1244–1286.

Sloutsky, V., Yim, H., Yao, X., & Dennis, S. (2017). An associative account of the development of word learning. *Cognitive Psychology*, *97*, 1–30.

Unger, L., Savic, O., & Sloutsky, V. (2020). Statistical regularities shape semantic organization throughout development. *Cognition*, *198*, 104190.

Unger, L., Vales, C., & Fisher, A. (2020). The role of co-occurrence statistics in developing semantic knowledge. *Cognitive Science*, *44*(9), 12894.

Wojcik, E. (2018). The development of lexical–semantic networks in infants and toddlers. *Child Development Perspectives*, *12*(1), 34–38.

Yim, H., Savic, O., Unger, L., Sloutsky, V., & Dennis, S. (2019). Can paradigmatic relations be learned implicitly? In A. Goel, C. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st annual conference of the cognitive science society* (p. 3389). Austin, TX: Cognitive Science Society.