Received: 27 July 2022 | Revised: 27 December 2022 | Accepted: 12 January 2023

DOI: 10.1111/desc.13373

Developmental Science WILEY

RESEARCH ARTICLE

No frills: Simple regularities in language can go a long way in the development of word knowledge

Layla Unger[1] | Hyungwook Yim[2] | Olivera Savic[1] | Simon Dennis[3] | Vladimir M. Sloutsky[1]

[1]Department of Psychology, The Ohio State University, Columbus, USA

[2]Department of Cognitive Sciences, Hanyang University, Seoul, Korea (the Republic of)

[3]Melbourne School of Psychological Sciences, The University of Melbourne, Melbourne, Australia

Correspondence
Layla Unger, Department of Psychology, The Ohio State University, Columbus, USA.
Email: unger.114@osu.edu

Hyungwook Yim, Department of Cognitive Sciences, Hanyang University, Seoul, Korea (the Republic of).
Email: hwyim@hanyang.ac.kr

Funding information
National Research Foundation of Korea, Grant/Award Number: RS-2022-00166828; National Institutes of Health, Grant/Award Numbers: P01HD080679, R01HD078545

Abstract

Recent years have seen a flourishing of Natural Language Processing models that can mimic many aspects of human language fluency. These models harness a simple, decades-old idea: It is possible to learn a lot about word meanings just from exposure to language, because words similar in meaning are used in language in similar ways. The successes of these models raise the intriguing possibility that exposure to word use in language also shapes the word knowledge that children amass during development. However, this possibility is strongly challenged by the fact that models use language input and learning mechanisms that may be unavailable to children. Across three studies, we found that unrealistically complex input and learning mechanisms are unnecessary. Instead, simple regularities of word use in children's language input that they have the capacity to learn can foster knowledge about word meanings. Thus, exposure to language may play a simple but powerful role in children's growing word knowledge.

KEYWORDS
distributional semantics, language acquisition, semantic development, semantic organization, word learning

Research Highlights

- Natural Language Processing (NLP) models can learn that words are similar in meaning from higher-order statistical regularities of word use.
- Unlike NLP models, infants and children may primarily learn only simple co-occurrences between words.
- We show that infants' and children's language input is rich in simple co-occurrence that can support learning similarities in meaning between words.
- We find that simple co-occurrences can explain infants' and children's knowledge that words are similar in meaning.

1 | INTRODUCTION

Starting in infancy, children begin to amass vocabularies that rapidly grow to contain tens of thousands of words. For decades, efforts to understand how children accomplish this feat have largely focused on how children map individual words to their meanings (e.g., Markman & Wachtel, 1988; Smith & Yu, 2008; Tomasello & Farrar, 1986). However, child word learning is all the more remarkable for going beyond individual mappings. Instead, children amass vocabularies that are *semantically organized* according to similarities in meaning between words (Arias-Trejo & Plunkett, 2013; Bergelson & Aslin, 2017). Such organization is a key aspect of word knowledge that plays fundamental roles in human language fluency, including in comprehension and word learning (Borovsky et al., 2016; Hutchison, 2003; Nation & Snowling, 1999; Neuman et al., 2011; Ouellette, 2006; Sahlgren, 2008). How do children acquire such richly organized word knowledge?

In principle, mere exposure to words in language is a viable source for semantically organized word knowledge because words with similar meanings tend to be used in language in similar ways (Harris, 1954; Landauer & Dumais, 1997; Miller & Charles, 1991; Rubenstein & Goodenough, 1965). Specifically, words similar in meaning such as "apple" and "banana" tend to be used in *similar contexts of other words* (henceforth *similar contexts*), as in "I'd like an apple for my snack" and "I'd like a banana for my snack". The simple idea that regularities of word use can be informative about word meanings is commonly referred to as the *distributional hypothesis* (Figure 1). In recent years, the potential power of this idea has been amply demonstrated by distributional semantics models that successfully harness regularities of word use to capture the systematic similarities in meaning that organize human word knowledge (e.g., Jones & Mewhort, 2007; Landauer & Dumais, 1997; Lund & Burgess, 1996; Melamud et al., 2016; Mikolov et al., 2013; Pennington et al., 2014). Given these successes, do regularities of word use represent a viable source for semantically organized word knowledge during *human development*?

This possibility provides a promising explanation of how mere exposure to language may play a powerful role in the semantically organized word knowledge that begins to develop in infancy (Borovsky & Elman, 2006; Ervin, 1961; Landauer & Dumais, 1997; Sloutsky et al., 2017). This explanation has the advantage of being simultaneously simple and broad in scope. The breadth of the distributional hypothesis comes from its ability to account for learning similarity in meaning between

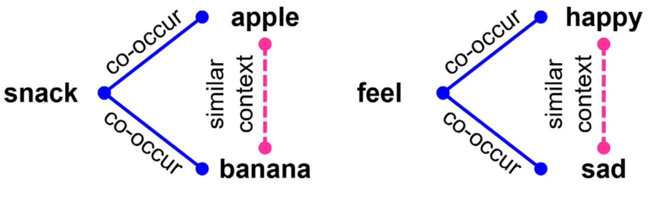

FIGURE 1 According to the distributional hypothesis, it is possible to learn that words such as "apple" and "banana" or "happy" and "sad" are similar in meaning because by tracking their co-occurrence with other words, and extrapolating that they occur similar contexts of other words.

any words encountered in language. This breadth contrasts with the popular proposal that similarity in meaning is learned based on the features that referents of words share, such as learning that "apple" and "banana" are similar because they both *are sweet* and *have skins* (Hills et al., 2009; Imai et al., 1994; McClelland & Rogers, 2003; Peters & Borovsky, 2019; Smith et al., 2002). Although such observable features may indeed provide one important source for semantically organized word knowledge, they are primarily useful for learning similarity in meaning between concrete nouns that denote objects with similar features. Therefore, in contrast with the distributional hypothesis, these accounts cannot explain learning similarity in meaning between many other words, such as "happy" and "sad" or "now" and "today". In addition to its breadth, the distributional hypothesis also has the simplicity of positing that learning similarity in meaning comes from exposure to language. Thus, this route avoids the need to assume that developing humans only learn similarities in meaning between words because they are pre-equipped with sophisticated abstract knowledge that there are systematic similarities in meaning between words (cf. Markman & Hutchinson, 1984; Xu & Tenenbaum, 2007).

This potential explanatory power has prompted some proposals that semantically organized word knowledge emerges from exposure to regularities of word use (e.g., Borovsky et al., 2012; Ervin, 1961; Fourtassi, 2020; Fremann & Lapata, 2015; Huebner & Willits, 2018; Landauer & Dumais, 1997; Sloutsky et al., 2017; Tamis-LeMonda et al., 2019). However, these proposals have overlooked the fact that for regularities of word use to provide a viable explanation of the *development* of semantically organized word knowledge, semantically-informative regularities that young children have the *capacity to learn* must be present in *their language input*. Moreover, these regularities must be able to account for the semantic organization that is acquired during human development.

As discussed below, prior demonstrations that semantic organization can emerge from exposure to regularities have not met these requirements. First, many such demonstrations have harnessed language input unavailable to young children, such as text from books or websites (Jones & Mewhort, 2007; Landauer & Dumais, 1997; Lund & Burgess, 1996; Mikolov et al., 2013; Pennington et al., 2014). Moreover, even the demonstrations that have harnessed language input available to children (Asr et al., 2016; Fourtassi et al., 2019; Frermann & Lapata, 2015; Huebner & Willits, 2018) have assumed that children possess learning mechanisms that may in fact remain immature for much of development. Specifically, such demonstrations have assumed that children have the capacity to learn that words are similar in meaning by tracking their co-occurrences with other words (solid blue lines in Figure 1) and extrapolating that they occur in similar contexts (dashed pink lines in Figure 1). For example, children are assumed to have the capacity to learn that "apple" and "banana" are similar based on hearing about having apples as a "snack" on one day, and bananas as a "snack" on another. Critically, this assumption is undermined by several lines of evidence reviewed below, which instead suggest that this capacity matures only gradually over the course of development. If the distributional hypothesis can only account for the development of semantically organized word knowledge by assuming access to input or learning mechanisms that young learners do not possess, then it is not a viable *developmental* account.

In what follows, we first highlight these challenges to the distributional hypothesis as a developmental account. We then present a series of studies that investigate whether the development of semantically organized word knowledge can emerge from regularities of word use given only the input and learning mechanisms known to be available during development.

1.1 | Challenge 1: State of the input

By four years of age, a child in a high SES family in the US has heard on average 45 million words, primarily from the people around them (e.g., Hart & Risley, 1995). In contrast, demonstrations that semantically organized word knowledge can emerge from regularities of word use commonly come from models trained on input that is both much more extensive and very different from the language children hear, such as text taken from books, websites, and newspapers (Jones & Mewhort, 2007; Landauer & Dumais, 1997; Lund & Burgess, 1996; Mikolov et al., 2013; Pennington et al., 2014). Such demonstrations provide little direct support for the viability of the distributional hypothesis as a developmental account. This viability instead depends on the presence of regularities in word use that can support learning similarity in meaning *in children's language input*.

Recent years have seen emerging evidence that the distributional hypothesis can overcome this challenge. For example, evaluations of children's language input such as in Tamis-LeMonda et al. (2019) provide evidence that words similar in meaning do indeed occur in similar contexts. Moreover, modeling studies conducted by Asr et al. (2016), Baroni et al. (2007), Fourtassi et al. (2019), Frermann and Lapata (2015), Huebner and Willits (2018) and Li et al. (2004) have found that distributional semantics models trained on corpora of recorded language input to children (MacWhinney, 2000) can learn similarities in meaning between words. These findings thus provide partial support for the distributional hypothesis as an account of the development of semantically organized word knowledge.

1.2 | Challenge 2: State of the learner

Even if young children receive language input that is rich in semantically-informative regularities of word use, do they have the capacity to exploit these regularities? Young learners are indeed sensitive to the simple regularities with which words *co-occur with each other*, such as the co-occurrence of "snack" and "apple" or "snack" and "banana" (see solid blue lines in Figure 1; Bannard & Matthews, 2008; Fisher, 2010; Fisher et al., 2011; Matlen et al., 2015; Unger, Savic, & Sloutsky, 2020; Unger, Vales, & Fisher, 2020; Wojcik & Saffran, 2015). Evidence from other domains suggests that this involves sensitivity not just to the frequency with which different inputs co-occur, but the regularity with which inputs co-occur more with each other than they do with other inputs (e.g., Aslin et al., 1998; Fiser & Aslin, 2002). However, efforts to cast the distributional hypothesis as a developmental account have gone beyond sensitivity to co-occurrence regularities. Instead, this work has assumed that a learner who can track such co-occurrences will also collaterally learn similarity between words such as "apple" and "banana" that occur in similar co-occurrence contexts, *even if they never co-occur with each other* (see dashed pink lines in Figure 1). This assumption is fundamental to distributional semantics models, which take a variety of forms but all possess mechanisms for learning similarity in meaning based on occurrence in similar contexts (for reviews of these mechanisms, see Lenci, 2018; Turney & Pantel, 2010; Willits et al., 2016).

The capacity to learn occurrence in similar contexts requires learning higher-order regularities that are not taken directly from the input itself, and instead are extrapolated across entirely different experiences, such as hearing a sentence about having apples as a snack on one day, and a sentence about bananas as a snack on another. Critically, the assumption that young learners possess this capacity may be unfounded.

One line of evidence undermining this assumption comes from studies showing that even adults struggle to learn that words are related based only on occurrence in similar contexts (e.g., Frigo & McDonald, 1998; Mintz, 2002; Ouyang et al., 2017). Instead, adults only learned that words that occurred in similar contexts were related when: (A) occurrence in similar contexts was correlated with additional cues, such as when many of the words that occur in similar contexts also contain the same speech sound, or (B) the contexts consisted of rigid, multi-word frames, in which the same fixed set of words occurred both before and after the words that occurred in the same context. Importantly, many of these studies focused on the problem of learning relations between words in the same *grammatical category* (such as *noun* or *verb*), which may indeed be marked by correlated cues or fixed multi-word frames in children's language input (Mintz, 2003; Monaghan et al., 2007). In contrast, there is as yet no evidence that these additional sources of information are available for learning that words are *similar in meaning*.

Evidence regarding *children's* ability to learn that words are related based on their occurrence in similar contexts is thin on the ground. This evidence primarily comes from a handful of studies with infants (Gerken et al., 2005; Lany & Saffran, 2010, 2011, 2013), in which learning only occurred when occurrence in similar contexts was correlated with additional cues. In contrast, one recent series of studies suggests that the capacity to learn that words are similar in meaning based *solely* on their occurrence in similar contexts develops very gradually (Savic et al., 2022). To study the development of this capacity, these studies simplified the challenge of learning the occurrence of words in similar contexts, so that one pair of words occurred in one context, and another pair of words occurred in a different context. Given these simple regularities of word use, adults successfully learned that words that occurred in similar contexts were similar in meaning. In contrast, this capacity was largely absent in 4-year-old children, and remained well below adult levels in 7–8-year-old children. Moreover, children's failure to learn occurrence in similar contexts remained even given multiple days of exposure, suggesting that this ability develops slowly even with extensive exposure and opportunities for consolidation.

This protracted developmental trajectory is corroborated by studies of learning analogous regularities in other domains. For example, several studies conducted by Bauer and colleagues (e.g., Bauer et al., 2020; Bauer & San Souci, 2010) have investigated children's capacity to integrate separate facts that are both connected to the same concept, such as integrating across the facts "dolphins live in groups called pods" and "dolphins communicate by clicking and squeaking" to learn that "pods communicate by clicking and squeaking". In 4-year-old children, evidence of integration is weak and largely dependent on explicit prompts during a subsequent test. As in Savic et al. (2022)'s studies with words, the capacity to integrate develops only gradually and remains below adult levels even in 9-year-old children. Similarly protracted developmental trajectories have been observed in studies of forming memories that span different experiences (Schlichting et al., 2021, 2017; Shing et al., 2019). Taken together, these lines of evidence dispute the capacity to learn occurrence in similar contexts assumed in the distributional hypothesis. By the same token, this evidence leaves the distributional hypothesis on very shaky footing as a viable developmental account.

1.3 | Present research

In principle, exposure to regularities of word use may provide a simple but powerful source for the development of semantically organized word knowledge. However, this possibility is challenged by strong limits on the input and learning mechanisms available during human development. Can semantically organized word knowledge emerge from the simple regularities of word use that young children can learn and are present in their language input?

To answer this question, we investigated the semantic information available from regularities of word use in language input to infants and children (MacWhinney, 2000). Critically, unlike prior evaluations of the distributional hypothesis as a developmental account (Asr et al., 2016; Fourtassi et al., 2019; Frermann & Lapata, 2015; Huebner & Willits, 2018), we focused on the simple regularity with which words reliably co-occur *with each other*. As noted above, the assumption that children can track these co-occurrences rests on a solid foundation of evidence (Bannard & Matthews, 2008; Fisher, 2010; Fisher et al., 2011; Matlen et al., 2015; Unger, Savic, & Sloutsky, 2020; Unger, Vales, & Fisher, 2020; Wojcik & Saffran, 2015). Moreover, preliminary evidence suggests that simple co-occurrences play an important role in early vocabulary growth (Flores et al., 2020). However, this evidence does not illuminate whether the capacity to track these simple co-occurrences can support the acquisition of *semantically organized* vocabularies. The current investigation builds upon this evidence to evaluate whether simple co-occurrences in children's language input can foster the development of semantically organized word knowledge.

We investigated this question in three studies that harnessed existing datasets. Study 1 provides a proof-of-principle by evaluating whether language input to infants and children is rich in simple co-occurrence regularities that can support learning that words similar in meaning (such as words for colors, emotions, or vehicles) are related. Study 2 builds on this proof-of-principle by investigating whether these simple co-occurrence regularities can account for the semantic organization that has been documented in early human development across several prior studies. Finally, Study 3 harnesses a dataset of word associations across childhood (Wojcik & Kandhadai, 2019) to investigate whether co-occurrence regularities can explain the *strength* of semantic associations between words across child development. Findings from all three studies provide robust evidence that even simple co-occurrence regularities in the language that infants and children hear can foster the development of semantically organized word knowledge.

2 | METHODS AND RESULTS

2.1 | All studies: Co-occurrence in language input to infants and children

All data and scripts used for the present studies are available on OSF: https://osf.io/5qajs/?view_only=2378e376a7a44985b97e6296b7fb2ad3.

All studies used a common measure of co-occurrence regularities between pairs of words in corpora of recorded and transcribed language input to infants and children, which we accessed from the CHILDES database (MacWhinney, 2000). Prior to measuring co-occurrence regularities, we first followed common corpus pre-processing procedures, including removing punctuation and standardizing morphological variants (e.g., singular and plural forms of nouns and tenses of verbs) to a single word form using the (Rinker, 2018). In addition, we manually identified words across all studies that were variants of a word – e.g., "bike" and "bicycle", "fridge" and "refrigerator" – and merged these into a common form. The pre-processed corpora contained approximately 3M words and 11,000 word types.

We then measured co-occurrence regularities between pairs of words in the corpora. The results presented here use a measure called a "t-score" (Evert, 2008); and we additionally replicated all analyses and found equivalent results using a Pointwise Mutual Information measure. These measures, as well as measures of raw co-occurrence frequencies, are included in the materials on OSF. To avoid confusion with the t-statistic from a Student's t-test, henceforth we refer to this score as a "co-score". Co-scores measure the degree to which pairs of words co-occur more frequently than would be expected by chance based on their individual frequencies. To measure the regularity with which a pair of words co-occur, co-scores use the following co-occurrence frequencies: (1) Co_1, the summed frequency of co-occurrence of word 1 with any other word; (2) Co_2, the summed frequency of co-occurrence of word 2 with any other word, (3) Co_{12}, the frequency of co-occurrence of word 1 with word 2, and (4) Co_{total}, the summed frequencies of co-occurrence of all pairs of words. These frequencies are first used to calculate the frequency of co-occurrence

TABLE 1 Example calculation of co-scores.

$word_1$	$word_2$	Co_1	Co_2	Co_{total}	E_{12}	Co_{12}	co-score
dog	cat	363878	284042	224042226	461.33	5016	64.31
dog	see	363878	2993292	224042226	4861.55	6048	15.26

between words 1 and 2 expected by chance, E_{12}:

$$E_{12} = \frac{Co_1\, Co_2}{Co_{total}} \tag{1}$$

Co-scores then compare the actual co-occurrence frequency of words 1 and 2 to the frequency expected by chance as follows:

$$co - score = \frac{Co_{12} - E_{12}}{\sqrt{Co_{12}}} \tag{2}$$

Table 1 provides an example of these calculations for the regularity with which “dog” co-occurs with “cat” and “see” in an 11-word window. The frequency with which “dog” co-occurs with “see” is somewhat higher than “cat”. However, because “see” is a frequent word that co-occurs with many other words, the frequency of its co-occurrence specifically with “dog” is only modestly higher than would be expected by chance. In contrast, the co-occurrence with “cat” is substantially higher than would be expected by chance.

We calculated co-scores for co-occurrence in sliding windows of 3, 7, and 11 words (which could span language input identified as different utterances in the CHILDES corpora). Each window size implements a constraint on how close together words must co-occur for a hypothetical child to learn that they are semantically related. The purpose of including multiple windows was to remain agnostic about the distance across which children can track co-occurrences (see e.g., Hills et al., 2010, for evidence that multiple window sizes may be relevant for capturing child word knowledge). We thus tested whether findings in each study were consistent across such constraints.

2.2 | Study 1

Study 1 was designed to provide a proof-of-principle by investigating whether language input to infants and children contains simple co-occurrence regularities that can support learning that words similar in meaning are related. We first identified words similar in meaning as words belonging to the same semantic category, such as words for colors, and then assessed whether words in the same semantic category tend to regularly co-occur.

2.2.1 | Semantic categories

We adapted semantic categories from the categories used in the MacArthur-Bates Communicative Index (MCDI) (Fenson et al., 2007; Frank et al., 2016), a widely-used measure of the words that infants and children know. The MCDI includes words from multiple grammatical classes (including nouns, adjectives, and verbs), divided into categories such as words for animals or clothes. However, some categories, such as “descriptive words”, are very broad and are not clearly semantic categories of words similar in meaning. We therefore modified some broad categories into more semantically coherent categories, such as “color words”. All modified categories were normed with adults to confirm that adults judged them to be semantically coherent categories of words similar in meaning (see Supplemental Materials). In addition, we supplemented the words in the MCDI with words belonging to these categories that occurred frequently (at least 100 times) in CHILDES corpora. To support the generalizability of findings, the selected 27 semantic categories came from multiple grammatical categories, including nouns (e.g., animals), adjectives (e.g., colors), and verbs (e.g., locomotion verbs).

2.2.2 | Co-occurrence regularities in semantic categories

Our analysis assessed the degree to which words in the same semantic category tend to reliably co-occur. First, we generated a “co-occurrence density” measure of reliable co-occurrence within a set of words. We calculated co-occurrence density as the proportion of word pairs within a set that had a co-score above a certain cutoff value. For analyses presented here, we used a cutoff of 1 standard deviation above the mean (see Supplemental Materials for similar results for a 2 standard deviation-cutoff). Co-occurrence density ranged from 0 to 1, such that higher values indicate that a higher proportion of words within a set reliably co-occur.

Next, we used a bootstrap approach to evaluate whether words in the same semantic category co-occur more reliably (i.e., have higher co-occurrence densities) than words from different categories. In this approach, we measured co-occurrence density within a semantic category, and compared it to co-occurrence densities for 1000 “matched” sets of words sampled from across categories. The matched sets for a semantic category contained the same number of words as the category, but sampled across categories so that the sampled words matched the category words in log frequency in CHILDES corpora. To match sampled words in log frequency, we calculated the log frequency of each word in the CHILDES corpus and divided these values into quantiles. For a given category, we then calculated the number of words in each quantile, and generated a log frequency-matched set of random words by sampling the same number of words in each quantile across categories. Finally, after sampling matched sets of words, we calculated the difference in co-occurrence density between a semantic category

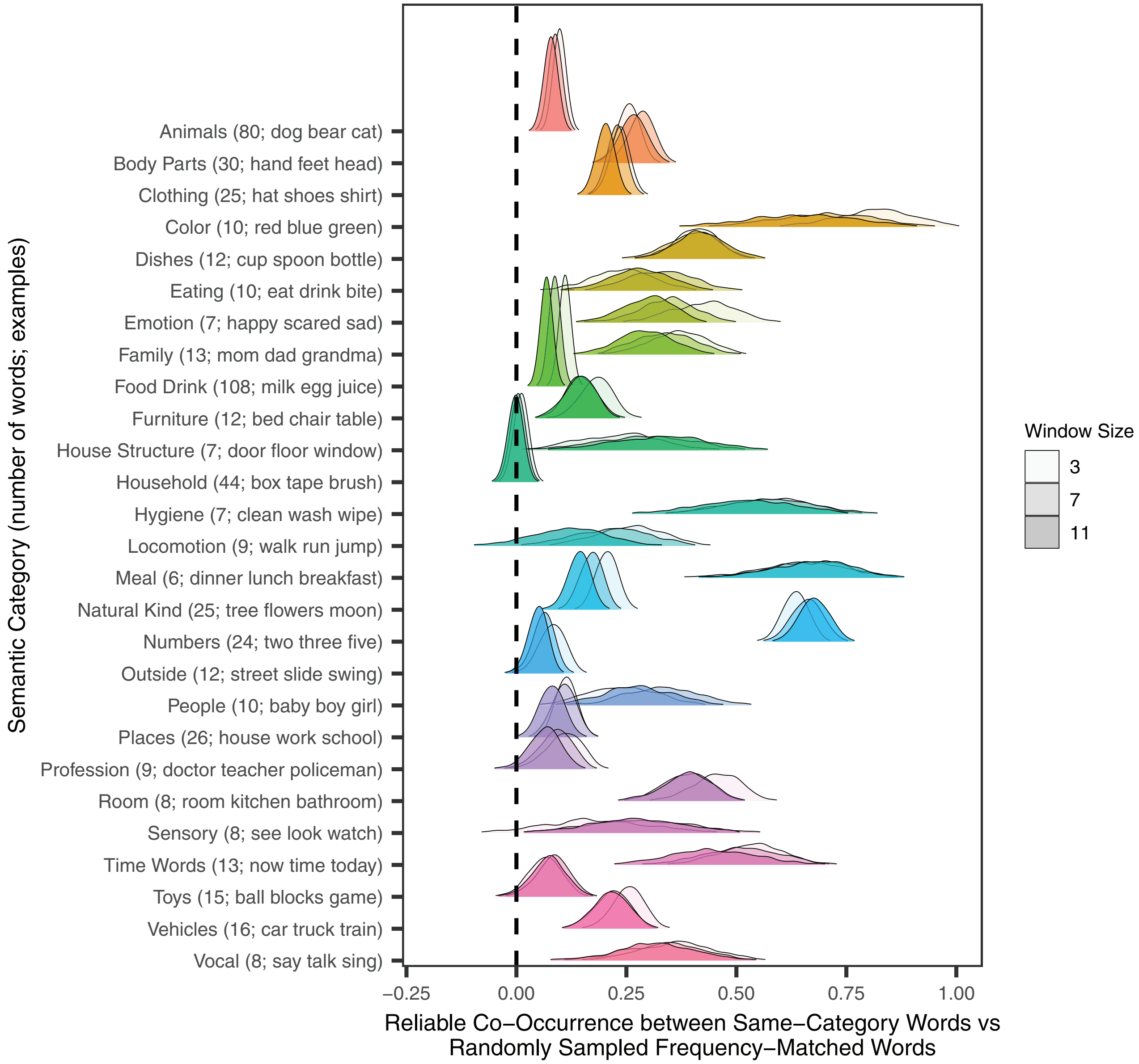

FIGURE 2 Each distribution depicts the distribution of differences in co-occurrence density between words in a semantic category, versus frequency-matched samples of words. Values greater than 0 indicate higher co-occurrence densities within a semantic category than frequency-matched samples of words. For the majority of categories and across window sizes, words from the same semantic category co-occur more reliably than sets of words matched in frequency. Thus, co-occurrence regularities capture similarity in meaning between words from the same semantic category.

and each of its matched sets, such that values greater than 0 indicate that words in the semantic category co-occur more reliably than words in different categories. We inferred that words in the same category tend to reliably co-occur when more than 95% of these differences were greater than 0.

Figure 2 shows distributions of difference values for each category and window size. These graphs show a consistent tendency for words from the same semantic category to reliably co-occur in children's language input. The few exceptions consisted of household items and locomotion verbs in a window of 11 words, and sensory verbs in a window of 3 words (see Supplemental Materials for all statistics and for results of a complementary analysis of the degree to which words tend to co-occur with words from the same category). Thus, across many semantic categories, even simple co-occurrence regularities in children's language input can support learning that words similar in meaning are related. For example, to learn that words for different vehicles are related, it is sufficient to be sensitive to the regularity with which these words co-occur with each other.

2.3 | Study 2

A number of prior studies have found that even as infants and toddlers are beginning to learn words, the word knowledge they acquire is becoming semantically organized (Arias-Trejo & Plunkett, 2013; Bergelson & Aslin, 2017; Chow et al., 2017; Delle Luche et al., 2014; Nguyen, 2007; Rämä et al., 2013; Sirri & Rämä, 2015; Styles & Plunkett, 2009; Willits et al., 2013). For example, Bergelson and Aslin (2017) provided evidence that words similar in meaning are linked even in the very early lexicons of 6-month-old infants. The results of Study 1 point to the possibility that this semantic organization may come from co-occurrence regularities in the language input to young learners. However, these results do not illuminate whether co-occurrence regularities can account for the semantic organization that has actually been recorded in early development. Indeed, many of the studies that have investigated early semantic organization have sought to eliminate this possibility. Specifically, these studies have controlled for "association strength", an indirect measure of co-occurrence based on the degree to which normative samples of participants (typically adults) tend to respond with one word when prompted with another in a free association task.

In Study 2, we therefore investigated whether co-occurrence regularities can account for the semantic organization that has been recorded in early human development. To conduct this investigation, we identified published studies that assessed early semantic organization based on infants' and toddlers' ability to differentiate between semantically related versus unrelated items. Ideally, we would then test whether the degree to which words reliably co-occur predicts the degree to which infants and toddlers know that they are semantically related (though it is worth noting that such a test would be complicated by the noisiness of infant and toddler data and the small sets of items used in most prior studies). However, the majority of published studies do not provide this item-level information. Therefore, we evaluated whether co-occurrence regularities could account for this differentiation by assessing whether semantically related items used in these studies also co-occurred more reliably than unrelated items.

To anticipate our findings, co-occurrence regularities consistently differentiated between semantically related and unrelated words across all studies. Moreover, this result transpired even for studies that were designed to control for co-occurrence by controlling for association strength.

2.3.1 | Selection of studies

For Study 2, we identified any prior study that investigated infants' and/or toddlers' ability to differentiate between semantically related versus unrelated items. Of these studies, we focused on studies that investigated semantic knowledge within a single language. In addition, we included only studies for which we were able to identify stimuli, either from the original paper or by contacting the authors (Arias-Trejo & Plunkett, 2013; Bergelson & Aslin, 2017; Chow et al., 2017; Delle Luche et al., 2014; Nguyen, 2007; Rämä et al., 2013; Sirri & Rämä, 2015; Styles & Plunkett, 2009; Willits et al., 2013). Stimuli for all studies are available on OSF.

2.3.2 | Analysis approach

Analyses compared co-occurrence between items in semantically related (e.g., "hand"–"foot") versus unrelated (e.g., "hand"–"frog") conditions. Because most prior studies used small sets of items, we used permutation tests to conduct this comparison. The implementation of this analysis varied according to the design of the prior study being evaluated, as described below.

2.3.3 | Design 1: Related versus unrelated conditions

Some studies assessed the degree to which participants perceived various items as related, and compared this measure in semantically related versus unrelated conditions. For example, Willits et al. (2013) compared the amount of time that 24-month-old infants spent listening to semantically related versus unrelated word pairs.

For these studies, we simply calculated the *t-statistic* (i.e., the test statistic from a Student's *t*-test) for the difference in co-occurrence between items in semantically related versus unrelated conditions. We then compared this value to *t*-statistics for permuted samples in which the condition labels were shuffled. We inferred that co-occurrence regularities differentiated between semantically related and unrelated conditions when the *t*-statistic for the true conditions was greater than at least 95% of the *t*-statistics for the shuffled conditions. Co-occurrence differentiated between semantically related and unrelated conditions across all studies (Figure 3a).

2.3.4 | Design 2: Preference for a related target

The second type of design involved presenting participants with one item as a prompt and assessing whether they showed some form of preference for a related target over one or more unrelated distractors. For example Chow et al. (2017) assessed whether hearing a word prime such as "boat" prompted 24–30– month-old participants to look more at a picture of a semantically related target item such as "train", versus unrelated distractor items such as "hat". For these studies, we calculated a co-occurrence-based preference for the related target. Specifically, we calculated the degree to which the prompt item co-occurred more with the related target versus the unrelated distractor(s) using the Luce choice rule:

$$Preference(target) = \frac{co - score_{prompt/target}}{co - score_{prompt/target} + co - score_{prompt/distractor(s)}} \tag{3}$$

We then calculated a *t*-statistic comparing this co-occurrence-based preference to chance (e.g., to .5 for studies with one target and one distractor). Finally, we compared this *t*-statistic to *t*-statistics for permuted samples in which the related target and unrelated distractor labels were shuffled within trials. In all studies, co-occurrence predicted a significant preference for the related target (Figure 3b).

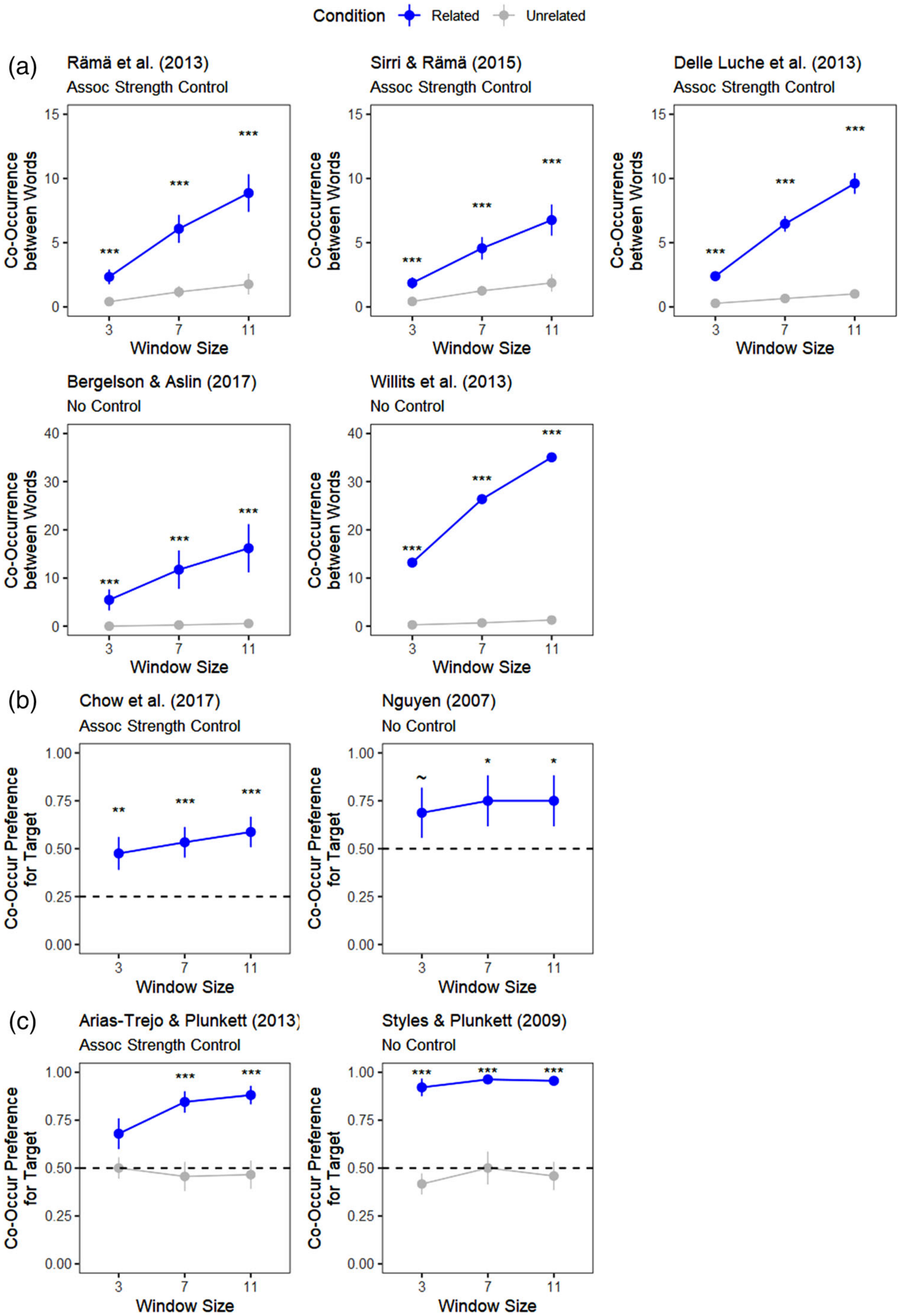

FIGURE 3 Co-occurrence regularities (co-scores) differentiate between semantically related and unrelated conditions. This finding was consistent across studies that used Design 1 (panel a), 2 (panel b) or 3 (panel c). Similarly, this finding transpired regardless of whether the study controlled for association strength (indicated in subtitles).

2.3.5 | Design 3: Effect of a related versus unrelated prime on preference for a named target

The final design comes from eye tracking studies in which participants saw a target and a distractor (e.g., sock and spoon), and heard the label of the target (e.g., "sock"). Critically, prior to the target label, participants heard a semantically related (e.g., "coat") or unrelated (e.g., "cat") prime word. Thus, semantic knowledge was assessed based on greater preferential looking at the labeled target versus the distractor when the label was preceded by a semantically related versus unrelated prime. This design is a combination of Designs 1 and 2, and our analysis combined our approaches for these designs. We first used the Luce choice rule equation above (Equation 3) to calculate a co-occurrence-based preference for the target in the semantically

Developmental Science WILEY

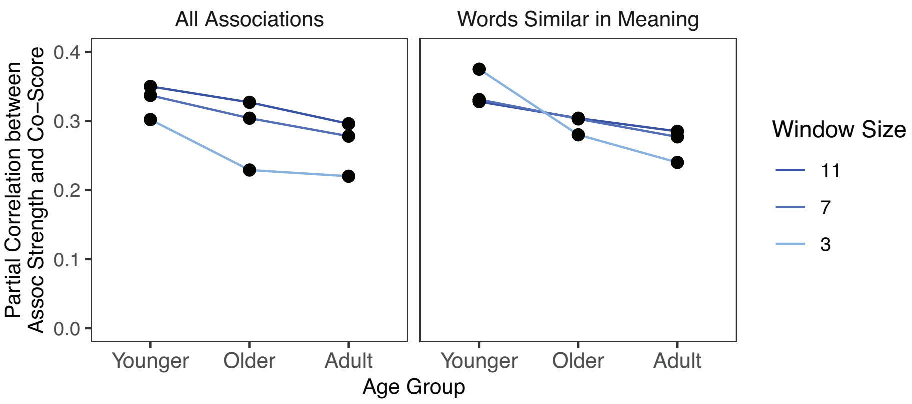

FIGURE 4 Relationship between association strength and co-scores across window sizes and age groups, depicted as partial correlations between these variables. Partial correlations were calculated while controlling for the log frequency of the response word, as in the analyses. As depicted in these graphs, co-scores predict association strength (recorded in the Wojcik & Kandhadai, 2019 dataset) across development, including associations between words similar in meaning.

related and in the unrelated prime conditions. Next, we calculated a *t*-statistic for the difference in preference between the semantically related and unrelated conditions, and compared this t-statistic to t-statistics calculated for permuted samples in which condition labels were shuffled (Figure 3c). This analysis revealed that co-occurrence predicted the greater preference for the target in semantically related versus unrelated conditions.

2.4 | Study 3

Study 2 provided evidence that co-occurrence regularities can explain prior evidence of semantically organized word knowledge early in development. Study 3 went a step further to investigate whether co-occurrence regularities can explain not only a binary distinction between semantically related and unrelated words, but can moreover account for the *strength* of semantic relations between words.

To accomplish this goal, Study 3 harnessed a dataset of spontaneous associations between words in children aged 3–7 years and adults (Wojcik & Kandhadai, 2019). This dataset was collected by prompting participants to respond to "cue" words with the first word that came to their minds. Cue words spanned multiple grammatical classes, including nouns, verbs, and adjectives. Free association data can be used to calculate the association strength between cue-response word pairs. Association strength is the proportion of participants who produced a response word to a cue word out of the total responses to the cue across all participants. Accordingly, we assessed whether co-occurrence predicts association strength between words. Following Wojcik and Kandhadai (2019), we assessed whether co-occurrence predicts association strength in younger children (3–5 years), older children (6–7 years),and adults. Importantly, we assessed this prediction both for all associations, and for just responses coded by the researchers as "paradigmatic" – that is, as a response similar in meaning to the cue, such as "stand" in response to "sit". To anticipate our results, we found that co-occurrence predicted association strength across development, including for words similar in meaning (Figure 4).

Our analyses used linear mixed effects models (Bates et al., 2015) with association strength as the outcome variable, co-occurrence and its interaction with age group as predictor variables, and a random intercept for cue word. To test whether the association strength of responses to cues was predicted by their co-occurrence with cues above and beyond their frequency, models also included the log frequency of the response word in CHILDES corpora and its interaction with age as predictors. For all window sizes, co-occurrence was a significant predictor (all $ps < .0001$), and did not interact with age (all $ps > .12$) (models including co-occurrence also outperformed baseline models including just response log frequency; all $ps < .0001$). Moreover, this result transpired when analyses were restricted to only responses coded as paradigmatic, and responses to noun, adjective, or verb cues (see Table 2 for full results). Thus, co-occurrence regularities robustly predict spontaneous associations between words from early in development onward, including associations between words similar in meaning.

2.4.1 | Co-occurrence regularities versus alternative predictors

As explained in the Introduction, the current research focused on the regularity with which words co-occur with each other (operationalized as co-scores) based on evidence that humans learn these regularities even early in development. To further evaluate the role of these regularities, we conducted two sets of analyses that compared co-scores as predictors of association strength to alternative predictors. The full results of these analyses are reported in Supplemental Materials. First, in line with the evidence reviewed above that children may struggle to learn occurrence in similar contexts compared to simple co-occurrence regularities, we found that co-scores were a more robust predictor of association strength than one of the multiple measures of occurrence

TABLE 2 Fixed effects in linear mixed effects models. Results are reported as χ^2 (*df*) and *p*-values.

Dataset	Window	Co-Occur	Freq	Co-Occur*Age	Freq*Age
Full	3	88.612(1), <**.0001**	14.444(1), <**.0001**	0.828(2), 0.661	195.846(2), <**.0001**
Full	7	141.566(1), <**.0001**	6.812(1), **0.009**	3.116(2), 0.211	189.271(2), <**.0001**
Full	11	162.744(1), <**.0001**	4.889(1), **0.027**	4.278(2), 0.118	187.985(2), <**.0001**
Paradigmatic	3	45.9(1), <**.0001**	12.024(1), **0.001**	0.715(2), 0.699	42.668(2), <**.0001**
Paradigmatic	7	56.625(1), <**.0001**	6.758(1), **0.009**	0.252(2), 0.882	36.44(2), <**.0001**
Paradigmatic	11	60.445(1), <**.0001**	5.501(1), **0.019**	0.413(2), 0.814	35.305(2), <**.0001**
Noun	3	65.356(1), <**.0001**	1.749(1), 0.186	1.115(2), 0.573	85.864(2), <**.0001**
Noun	7	87.415(1), <**.0001**	0.392(1), 0.531	1.95(2), 0.377	87.47(2), <**.0001**
Noun	11	93.904(1), <**.0001**	0.192(1), 0.662	2.567(2), 0.277	88.687(2), <**.0001**
Verb	3	6.19(1), **0.013**	7.69(1), **0.006**	3.192(2), 0.203	93.305(2), <**.0001**
Verb	7	14.219(1), <**.0001**	5.001(1), **0.025**	2.507(2), 0.286	90.621(2), <**.0001**
Verb	11	17.142(1), <**.0001**	4.168(1), **0.041**	2.624(2), 0.269	89.684(2), <**.0001**
Adjective	3	32.445(1), <**.0001**	1.08(1), 0.299	1.802(2), 0.406	38.062(2), <**.0001**
Adjective	7	50.07(1), <**.0001**	0.169(1), 0.681	4.731(2), 0.094	35.772(2), <**.0001**
Adjective	11	55.182(1), <**.0001**	0.073(1), 0.787	5.661(2), 0.059	35.53(2), <**.0001**

in similar contexts commonly used in distributional semantics models (cosine similarity). Second, in line with evidence that humans track the regularity with which inputs co-occur *with each other* rather than mere co-occurrence frequencies, we found that co-scores were also a better predictor than raw co-occurrence frequencies. These findings further support the viability of co-occurrence regularities as a source for semantically organized word knowledge.

3 | GENERAL DISCUSSION

As children learn words, the word knowledge they acquire becomes semantically organized. Regularities of word use provide a promising potential source for such organized word knowledge. However, prior demonstrations that semantically organized word knowledge can emerge from regularities of use (Asr et al., 2016; Fourtassi, 2020; Fourtassi et al., 2019; Frermann & Lapata, 2015; Huebner & Willits, 2018; Jones & Mewhort, 2007; Landauer & Dumais, 1997; Lund & Burgess, 1996) have harnessed input or learning mechanisms that may be unavailable to young children. The present studies therefore provide crucial new evidence that semantic organization can emerge from even simple co-occurrence regularities that infants and children can learn and that are present in their language input.

Study 1 demonstrated that words similar in meaning (such as words for animals) tend to reliably co-occur in infants' and children's' language input. Study 2 showed that these regularities are not only available in language input, but can account for the semantic organization of word knowledge observed in early development. Finally, in Study 3, we found that co-occurrence regularities can account for the strength of semantic associations between words throughout development. Together, these findings highlight how the development of semantically organized word knowledge may emerge from young children's everyday experiences with language.

3.1 | Synergies with other sources of word knowledge

The present studies focused on the development of word knowledge that captures similarities in meaning. This is a vital aspect of word knowledge because it turns what would otherwise be a jumble of random words into an organized mental lexicon. Accordingly, word knowledge that is organized according to similarity in meaning plays vital roles in language processes such as comprehension (Landi & Perfetti, 2007; Nation & Snowling, 1999) and learning new words (Borovsky et al., 2016; Neuman & Dwyer, 2011; Neuman et al., 2011; Sloutsky et al., 2017). The present findings provide evidence that children's early emerging abilities to track simple co-occurrence regularities are a viable source for developing semantically organized word knowledge.

However, children develop in an environment that is rich in linguistic and perceptual information. Synergies between co-occurrence and other sources of input may be vital in the development of semantic knowledge about words. One such synergy is that co-occurrence between words may provide an important route for *generalizing* knowledge about word referents, such as generalizing information learned about the referents of the word "apple" to the word "banana". Indeed, there is evidence that young children generalize what they learn about the referent of one word to the referent of another with which it co-occurs (Fisher, 2010; Fisher et al., 2011; Matlen et al., 2015; Ngo et al., 2021).

Another key synergy between co-occurrence and other sources may be to provide converging cues to semantic relatedness. For example, as noted in the Introduction, perceptual similarities between referents of concrete words (particularly nouns) can also be informative about similarities in meaning (Hills et al., 2009; Peters & Borovsky, 2019). Moreover, there is evidence that children link words that denote perceptually similar referents (Wojcik & Saffran, 2013). In addition, although the relationship between the sound of a word and its meaning is typically arbitrary, languages also contain instances in which words that sound similar are similar in meaning (e.g., the "gl" onset common in words related to vision or light in English, such as "gleem", "glimmer" and "glance") (Bergen, 2004; Dingemanse et al., 2015). Therefore, word co-occurrence may complement and reinforce other sources of word learning to build rich bodies of semantic word knowledge.

3.2 | Development of learning regularities of word use

A key motivation for the present studies consisted of evidence that even infants and young children can learn simple co-occurrence regularities, whereas regularities of occurrence in similar contexts that play a key role in the distributional hypothesis may be difficult to learn even for adults, and are particularly challenging for young children. However, as noted in the Introduction, evidence that can shed light on the development of children's sensitivity to regularities of word use is relatively limited.

For example, semantically organized word knowledge emerges over days, months, and years of accumulating language exposure. Moreover, influential accounts of the acquisition of semantically organized knowledge point to such extensive exposure and opportunities for memory consolidation as critical facets of this process (e.g., Kumaran et al., 2016). In contrast, studies to date have primarily assessed learning of statistical regularities given a single session of exposure (though see Savic et al., 2022, Study 2 for a recent exception). Therefore, future research into the development of learning statistical regularities of word use over extended periods of time could provide greater insight into the regularities that contribute to the development of semantically organized word knowledge.

Another vital but largely unexplored question is whether and how children integrate multiple cues to similarity in meaning across development. As noted in the preceding section, co-occurrence may complement other cues to similarity in meaning available for some words, such as the perceptual similarity of their referents or similarity in sound of the words themselves. Thus, future research could investigate whether children capitalize on this convergence when it is available, or flexibly utilize different cues when they do not converge. Moreover, to build beyond the evidence that children struggle to learn occurrence in similar contexts, future research could explore whether children nevertheless use this regularity to learn similarity in meaning when it is supported by other cues.

In addition, children's real-world language input contains more variability than the language input analyzed in the present study, given that the corpora were processed to merge morphological variants of words into a single word form. This processing implements an assumption that learned co-occurrence regularities can generalize across word forms: for example, that hearing the words "dogs" and "cats" co-occur contributes to learning a semantic association between "dog" and "cat". However, morphological learning typically unfolds over the first few years of life (Cazden, 1968; Marquis & Shi, 2012; Tomasello & Olguin, 1993). Thus, to shed further light onto how regularities of word use shape the emergence of semantically organized word knowledge, future research could investigate the development of generalizing statistical regularities of word use across word forms.

4 | CONCLUSION

In principle, regularities of word use in language may provide a rich source of information about word meanings. The present findings reveal that even within strong developmental constraints on input and learning mechanisms, these regularities can play a powerful role in building semantic knowledge about words during development. Therefore, exposure to regularities of word use provides a viable developmental account of how young children's burgeoning word knowledge becomes semantically organized.

ACKNOWLEDGEMENTS

This work was supported by the National Research Foundation of Korea funded by the Ministry of Science and ICT (No. RS-2022-00166828) awarded to Hyungwook Yim and by National Institutes of Health Grants R01HD078545 and P01HD080679 awarded to Vladimir M. Sloutsky.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest with respect to their authorship or the publication of this article.

DATA AVAILABILITY STATEMENT

Data and scripts have been made available via the Open Science Framework at https://osf.io/5qajs.

ETHICS STATEMENT

All activities involved in this research were approved by the Institutional Review Board at The Ohio State University.

ORCID

Layla Unger https://orcid.org/0000-0002-8888-1265
Olivera Savic https://orcid.org/0000-0002-7014-9017

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.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*(4), 321–324.

Asr, F. T., Willits, J. A., & Jones, M. N. (2016). Comparing predictive and co-occurrence based models of lexical semantics trained on child-directed speech. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society* (pp. 1092–1097).

Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning: The effect of familiarity on children's repetition of four-word combinations. *Psychological Science*, *19*(3), 241–248.

Baroni, M., Lenci, A., & Onnis, L. (2007). Isa meets lara: An incremental word space model for cognitively plausible simulations of semantic learning. In *Proceedings of the Workshop on Cognitive Aspects of Computational Language Acquisition*, (pp. 49–56).

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Bauer, P. J., Cronin-Golomb, L. M., Porter, B. M., Jaganjac, A., & Miller, H. E. (2020). Integration of memory content in adults and children: Developmental differences in task conditions and functional consequences. *Journal of Experimental Psychology: General*, *150*, 1259–1278.

Bauer, P. J., & San Souci, P. (2010). Going beyond the facts: Young children extend knowledge by integrating episodes. *Journal of Experimental Child Psychology*, *107*(4), 452–465.

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

Bergen, B. K. (2004). The psychological reality of phonaesthemes. *Language*, *80*(2), 290–311.

Borovsky, A., Ellis, E. M., Evans, J. L., & Elman, J. L. (2016). Semantic structure in vocabulary knowledge interacts with lexical and sentence processing in infancy. *Child Development*, *87*(6), 1893–1908.

Borovsky, A., & Elman, J. (2006). Language input and semantic categories: A relation between cognition and early word learning. *Journal of Child Language*, *33*(4), 759–790.

Borovsky, A., Elman, J. L., & Kutas, M. (2012). Once is enough: N400 indexes semantic integration of novel word meanings from a single exposure in context. *Language Learning and Development*, *8*(3), 278–302.

Cazden, C. B. (1968). The acquisition of noun and verb inflections. *Child Development* (pp. 433–448).

Chow, J., Davies, A. 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.

Delle Luche, C., Durrant, S., Floccia, C., & Plunkett, K. (2014). Implicit meaning in 18-month-old toddlers. *Developmental Science*, *17*(6), 948–955.

Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H., & Monaghan, P. (2015). Arbitrariness, iconicity, and systematicity in language. *Trends in Cognitive Sciences*, *19*(10), 603–615.

Ervin, S. M. (1961). Changes with age in the verbal determinants of word-association. *The American Journal of Psychology*, *74*(3), 361–372.

Evert, S. (2008). Corpora and collocations. *Corpus linguistics. An international handbook*, *2*, 1212–1248.

Fenson, L., Marchman, V. A., Thal, D. J., Dale, P. S., Reznick, J. S., & Bates, E. (2007). *MacArthur-Bates Communicative Development Inventories*. Paul H. Brookes Publishing.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences*, *99*(24), 15822–15826.

Fisher, A. V. (2010). What's in the name? Or how rocks and stones are different from bunnies and rabbits. *Journal of Experimental Child Psychology*, *105*(3), 198–212.

Fisher, A. V., Matlen, B. J., & Godwin, K. E. (2011). Semantic similarity of labels and inductive generalization: Taking a second look. *Cognition*, *118*(3), 432–438.

Flores, A. Z., Montag, J., & Willits, J. (2020). Using known words to learn more words: A distributional analysis of child vocabulary development. *arXiv preprint arXiv:2009.06810*.

Fourtassi, A. (2020). Word co-occurrence in child-directed speech predicts children's free word associations. In *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics* (pp. 49–53).

Fourtassi, A., Scheinfeld, I., & Frank, M. C. (2019). The development of abstract concepts in children's early lexical networks. In *Proceedings of the workshop on cognitive modeling and computational linguistics* (pp. 129–133).

Frank, M. C., Braginsky, M., Yurovsky, D., & Marchman, V. A. (2016). Wordbank: An open repository for developmental vocabulary data. *Journal of Child Language*, *44*, 677–694.

Fremann, L., & Lapata, M. (2015). Incremental bayesian category learning from natural language. *Cognitive Science*, *40*, 1333–1381.

Frigo, L., & McDonald, J. L. (1998). Properties of phonological markers that affect the acquisition of gender-like subclasses. *Journal of Memory and Language*, *39*(2), 218–245.

Gerken, L., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, *32*(2), 249–268.

Harris, Z. S. (1954). Distributional structure. *Word*, *10*(2-3), 146–162.

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.

Hills, T. T., Maouene, J., Riordan, B., & Smith, L. B. (2010). The associative structure of language: Contextual diversity in early word learning. *Journal of Memory and Language*, *63*(3), 259–273.

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. B. (2009). Categorical structure among shared features in networks of early-learned nouns. *Cognition*, *112*(3), 381–396.

Huebner, P. A., & Willits, J. A. (2018). Structured semantic knowledge can emerge automatically from predicting word sequences in child-directed speech. *Frontiers in Psychology*, *9*.

Hutchison, K. A. (2003). Is semantic priming due to association strength or feature overlap? A microanalytic review. *Psychonomic Bulletin & Review*, *10*(4), 785–813.

Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development*, *9*(1), 45–75.

Jones, M. N., & Mewhort, D. J. (2007). Representing word meaning and order information in a composite holographic lexicon. *Psychological Review*, *114*(1), 1–37.

Kumaran, D., Hassabis, D., & McClelland, J. L. (2016). What learning systems do intelligent agents need? Complementary learning systems theory updated. *Trends in Cognitive Sciences*, *20*(7), 512–534.

Landauer, T. K., & Dumais, S. T. (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.

Landi, N., & Perfetti, C. A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and Language*, *102*(1), 30–45.

Lany, J., & Saffran, J. R. (2010). From statistics to meaning: Infants' acquisition of lexical categories. *Psychological Science*, *21*(2), 284–291.

Lany, J., & Saffran, J. R. (2011). Interactions between statistical and semantic information in infant language development. *Developmental Science*, *14*(5), 1207–1219.

Lany, J., & Saffran, J. (2013). Statistical learning mechanisms in infancy. *Comprehensive Developmental Neuroscience: Neural Circuit Development and Function in the Brain*, *3*, 231–248.

Lenci, A. (2018). Distributional models of word meaning. *Annual Review of Linguistics*, *4*, 151–171.

Li, P., Farkas, I., & MacWhinney, B. (2004). Early lexical development in a self-organizing neural network. *Neural Networks*, *17*(8-9), 1345–1362.

Lund, K., & Burgess, C. (1996). Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, *28*(2), 203–208.

MacWhinney, B. (2000). *The CHILDES project: The database* (vol. 2). Psychology Press.

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

Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology*, *20*(2), 121–157.

Marquis, A., & Shi, R. (2012). Initial morphological learning in preverbal infants. *Cognition*, *122*(1), 61–66.

Matlen, B. J., Fisher, A. V., & Godwin, K. E. (2015). The influence of label co-occurrence and semantic similarity on children's inductive generalization. *Frontiers in Psychology*, *6*, 1146.

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

Melamud, O., Goldberger, J., & Dagan, I. (2016). context2vec: Learning generic context embedding with bidirectional LSTM. In *Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning* (pp. 51–61).

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*.

Miller, G. A., & Charles, W. G. (1991). Contextual correlates of semantic similarity. *Language and Cognitive Processes*, *6*(1), 1–28.

Mintz, T. H. (2002). Category induction from distributional cues in an artificial language. *Memory & Cognition*, *30*(5), 678–686.

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, *90*(1), 91–117.

Monaghan, P., Christiansen, M. H., & Chater, N. (2007). The phonological-distributional coherence hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology*, *55*(4), 259–305.

Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition*, *70*(1), B1–B13.

Neuman, S. B., & Dwyer, J. (2011). Developing vocabulary and conceptual knowledge for low-income preschoolers: A design experiment. *Journal of Literacy Research*, *43*(2), 103–129.

Neuman, S. B., Newman, E. H., & Dwyer, J. (2011). Educational effects of a vocabulary intervention on preschoolers' word knowledge and conceptual development: A cluster-randomized trial. *Reading Research Quarterly*, *46*(3), 249–272.

Ngo, C. T., Benear, S. L., Popal, H., Olson, I. R., & Newcombe, N. S. (2021). Contingency of semantic generalization on episodic specificity varies across development. *Current Biology*, *31*(12), 2690–2697.

Nguyen, S. P. (2007). Cross-classification and category representation in children's concepts. *Developmental Psychology*, *43*(3), 719–731.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, *98*(3), 554.

Ouyang, L., Boroditsky, L., & Frank, M. C. (2017). Semantic coherence facilitates distributional learning. *Cognitive Science*, *41*, 855–884.

Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532–1543).

Peters, R., & Borovsky, A. (2019). Modeling early lexico-semantic network development: Perceptual features matter most. *Journal of Experimental Psychology: General*, *148*(4), 763.

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.

Rinker, T. W. (2018). *textstem: Tools for stemming and lemmatizing text*. Buffalo, New York. Version 0.1.4. http://github.com/trinker/textstem

Rubenstein, H., & Goodenough, J. B. (1965). Contextual correlates of synonymy. *Communications of the ACM*, *8*(10), 627–633.

Sahlgren, M. (2008). The distributional hypothesis. *Italian Journal of Disability Studies*, *20*, 33–53.

Savic, O., Unger, L., & Sloutsky, V. M. (2022). Experience and maturation: The contribution of co-occurrence regularities in language to the development of semantic organization. *Child Development*, (p. Available from: https://doi.org/10.31234/osf.io/5tpf2)

Schlichting, M. L., Guarino, K. F., Roome, H. E., & Preston, A. R. (2021). Developmental differences in memory reactivation relate to encoding and inference in the human brain. *Nature Human Behaviour*, *6*, 415–428.

Schlichting, M. L., Guarino, K. F., Schapiro, A. C., Turk-Browne, N. B., & Preston, A. R. (2017). Hippocampal structure predicts statistical learning and associative inference abilities during development. *Journal of Cognitive Neuroscience*, *29*(1), 37–51.

Shing, Y. L., Finke, C., Hoffmann, M., Pajkert, A., Heekeren, H. R., & Ploner, C. J. (2019). Integrating across memory episodes: Developmental trends. *PLoS ONE*, *14*(4), e0215848.

Sirri, L., & Rämä, P. (2015). Cognitive and neural mechanisms underlying semantic priming during language acquisition. *Journal of Neurolinguistics*, *35*, 1–12.

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

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, *13*(1), 13–19.

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

Styles, S. J., & Plunkett, K. (2009). How do infants build a semantic system? *Language and Cognition*, *1*(1), 1–24.

Tamis-LeMonda, C. S., Custode, S., Kuchirko, Y., Escobar, K., & Lo, T. (2019). Routine language: Speech directed to infants during home activities. *Child Development*, *90*(6), 2135–2152.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. *Child Development*, *57*(6), 1454–1463.

Tomasello, M., & Olguin, R. (1993). Twenty-three-month-old children have a grammatical category of noun. *Cognitive Development*, *8*(4), 451–464.

Turney, P. D., & Pantel, P. (2010). From frequency to meaning: Vector space models of semantics. *Journal of artificial intelligence research*, *37*, 141–188.

Unger, L., Savic, O., & Sloutsky, V. M. (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*, e12894.

Willits, J. A., Jones, M. N., & Landy, D. (2016). Learning that numbers are the same, while learning that they are different. In *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 1799–1804). Philadelphia, PA.

Willits, J. A., Wojcik, E. H., Seidenberg, M. S., & Saffran, J. R. (2013). Toddlers activate lexical semantic knowledge in the absence of visual referents: Evidence from auditory priming. *Infancy*, *18*(6), 1053–1075.

Wojcik, E. H., & Kandhadai, P. (2019). Paradigmatic associations and individual variability in early lexical–semantic networks: Evidence from a free association task. *Developmental Psychology*, *56*, 53–69.

Wojcik, E. H., & Saffran, J. R. (2013). The ontogeny of lexical networks: Toddlers encode the relationships among referents when learning novel words. *Psychological Science*, *24*(10), 1898–1905.

Wojcik, E. H., & Saffran, J. R. (2015). Toddlers encode similarities among novel words from meaningful sentences. *Cognition*, *138*, 10–20.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as bayesian inference. *Psychological Review*, *114*(2), 245–272.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Unger, L., Yim, H., Savic, O., Dennis, S., & Sloutsky, V. M. (2023). No frills: Simple regularities in language can go a long way in the development of word knowledge. *Developmental Science*, e13373. https://doi.org/10.1111/desc.13373