Statistical Regularities Shape Semantic Organization throughout Development

Layla Unger, **Olivera Savic**, **Vladimir M. Sloutsky**

Department of Psychology, Ohio State University, Columbus OH

Abstract

Our knowledge about the world is represented not merely as a collection of concepts, but as an organized lexico-semantic network in which concepts can be linked by relations, such as "taxonomic" relations between members of the same stable category (e.g., *cat* and *sheep*), or association between entities that occur together or in the same context (e.g., *sock* and *foot*). To date, accounts of the origins of semantic organization have largely overlooked how sensitivity to statistical regularities ubiquitous in the environment may play a powerful role in shaping semantic development. The goal of the present research was to investigate how associations in the form of statistical regularities with which labels for concepts co-occur in language (e.g., *sock* and *foot*) and taxonomic relatedness (e.g., *sock* and *pajamas*) shape semantic organization of 4–5-year-olds and adults. To examine these aspects of semantic organization across development, we conducted three experiments examining effects of co-occurrence and taxonomic relatedness on cued recall (Experiment 1), word-picture matching (Experiment 2), and looking dynamics in a Visual World paradigm (Experiment 3). Taken together, the results of the three experiments provide evidence that co-occurrence-based links between concepts manifest in semantic organization from early childhood onward, and are increasingly supplemented by taxonomic links. We discuss these findings in relation to theories of semantic development.

Keywords

semantic memory; semantic development; semantic organization; knowledge organization; statistical regularities; cognitive development

Our knowledge about the world is fundamental to many cognitive feats we accomplish on an everyday basis, including applying what we know to new situations (e.g., expecting new

home appliances to be powered by electricity), retrieving previously acquired knowledge from memory, and incorporating new information into our existing body of knowledge (Bower, Clark, Lesgold, & Winzenz, 1969; Heit, 2000; Jimura, Hirose, Wada et al., 2016). These feats are possible because our knowledge about the world is not a collection of isolated facts, but rather an interconnected lexico-semantic network of related concepts (Cree & Armstrong, 2012; Jones, Willits, & Dennis, 2015; McClelland & Rogers, 2003). For example, our knowledge about dogs is often connected to our knowledge of other similar animals (such as cats), as well as to our knowledge about the items with which dogs are associated in the environment, such as leashes, human owners, and doghouses.

Although the fact that our concepts are organized is hardly controversial (Howard & Howard, 1977; McClelland & Rogers, 2003; Ross & Murphy, 1999; Storm, 1980), the processes that drive the development of semantic organization are a topic of considerable debate. To date, this debate has primarily focused on how connections between concepts from the same stable, "taxonomic" category (e.g., *animals, foods*) are formed, in spite of the fact that taxonomic relatedness may be difficult to observe because members of the same taxonomic category do not necessarily look similar, or occur together[1].

Some proposals suggest that the development of knowledge organization starts with easy to observe relations, which then both bootstrap and are overwritten by knowledge of taxonomic relations (Inhelder & Piaget, 1964; Lucariello, Kyratzis, & Nelson, 1992). Alternately, other proposals suggest that we are endowed with early-emerging biases towards learning taxonomic relations, such as a belief that items that are referred to by the same label (e.g., "animal") belong to the same taxonomic category (Fulkerson & Waxman, 2007; Gelman & Coley, 1990; Gelman & Markman, 1986).

The goal of the current research is to investigate yet another possibility: That easy to observe relations – specifically, co-occurrence – play a fundamental role in knowledge organization from early in development through adulthood. Specifically, co-occurrence may directly foster the formation of associations between corresponding concepts, thus linking items such as *spaghetti, fork, plate* and *napkin*. Moreover, co-occurrence may also indirectly foster the formation of links between concepts that share patterns of co-occurrence, which are often taxonomically related, such as *spaghetti* and *pie* (which both co-occur with *fork, plate* and *napkin*) (see Jones et al., 2015, for a review of mechanistic models of forming relations from co-occurrence regularities). Additionally, because co-occurrence regularities can be experienced directly, links between concepts based on these regularities may manifest in semantic organization beginning early in development. In contrast, shared patterns of co-occurrence must be integrated across separate instances of co-occurrence. Therefore, learning taxonomic relations from shared patterns of co-occurrence may require more time, such that taxonomic relations may emerge more gradually in the course of development (Sloutsky, Yim, Yao, & Dennis, 2017). Importantly, according to this view, taxonomic relations supplement rather than replace co-occurrence relations.

---

[1]Here and in most prior theoretical accounts and research into semantic organization, the challenge is to explain the development of semantic links between *different* concepts (e.g., between *dog* and other animals), not just the emergence of basic-level categories (e.g., *dog*), for which many cues (e.g., shared labels and visual similarity) are readily available.

In what follows, we first review extant theoretical accounts that have focused on explaining the emergence of taxonomic relations in knowledge organization. We then highlight key findings indicative of a role for co-occurrence that these accounts fail to capture, and an alternate perspective in which co-occurrence contributes substantially to semantic development. Finally, we present three experiments designed to examine the presence of co-occurrence-based links and taxonomic links in semantic organization across development.

Traditional Accounts of the Emergence of Semantic Organization

Most extant accounts of the development of semantic organization have focused on how semantic knowledge becomes organized according to the membership of concepts in stable, taxonomic categories, such as *foods*. According to some of these accounts, referred to here as *restructuring* accounts, taxonomic relations are the endpoint of maturational/learning processes that unfold across development. These perspectives fall within a broader class of cognitive development accounts in which knowledge or abilities present early in development are supplanted by their successors (e.g., Carey, 1985).

Critical to these accounts is the idea that the order in which concepts and their relations are acquired is dictated by how observable they are. For example, it is easy to observe that dogs reliably co-occur with leashes or bones. In contrast, the membership of separate concepts in the same taxonomic category is often more difficult (if not impossible) to observe: For example, animals can be quite different from one another, and they do not necessarily appear together or in the same environment or context. Restructuring accounts propose that early organization is shaped by information readily available in the environment, and that over the course of development, taxonomic knowledge overwrites (or replaces) this form of (more rudimentary) organization.

An early restructuring account was proposed by Inhelder and Piaget (1964). According to this account, the transition towards taxonomic organization is driven by experiences that highlight the inadequacy of earlier modes of organization (although the mechanisms by which this transition occurs were not clearly explicated).

Another, more clearly specified restructuring account is Nelson and Lucariello's (e.g., Lucariello et al., 1992) slot-filler account. This account highlights the fact that both language and experience in the world contain regularities in which members of the same taxonomic category often play the same role in the same context. For example, some members of the taxonomic category of *foods*, such as pancakes, eggs, and bacon, reliably play the role of *being eaten* in a *breakfast* context. According to the slot-filler account, young children are sensitive to these regularities, such that semantic knowledge first becomes organized into contextually-constrained taxonomic groups. With development, contextually-independent taxonomic organization emerges as children come to abstract across these constrained taxonomic groups and recognize when entities play the same role in different contexts, such as foods *being eaten* in breakfast, lunch, and dinner contexts.

According to another set of accounts, referred to here as *taxonomic bias* accounts, taxonomic relations predominate semantic organization from early in development. This

type of account acknowledges that taxonomic relatedness is not directly apparent from environmental input, but posits that it is nonetheless apprehended early in development due to early-emerging (possibly innate) biases towards learning which entities are members of the same taxonomic category. For example, members of the same basic-level taxonomic category are often referred to by the same label, such as "bird". Accordingly, taxonomic bias accounts propose that we are endowed with early-emerging beliefs that entities in the world belong to taxonomic categories, and that entities that are referred to using the same label belong to the same category (Fulkerson & Waxman, 2007; Gelman & Coley, 1990; Gelman & Markman, 1986). Therefore, in these accounts, learning that starts early in development consists of using these labels to anchor the organization of semantic knowledge into basic-level taxonomic categories. However, it is less clear how these biases support the formation of relations *between* these categories that are crucial to semantic organization, such as relations between *birds* and other animal categories. In addition, a role for other types of input, such as the regularity with which entities co-occur, is not specified in these accounts.

A final type of account reviewed here, which we refer to as *featural learning*, posits that the development of semantic organization is driven by detecting clusters of reliably correlated features that are often associated with taxonomic categories (Rosch, 1975, 1978). For example, membership in the category of *birds* is associated with possessing *wings*, *feathers*, and a *beak*. Featural learning accounts propose learning mechanisms that are sensitive to these correlations, and that the operation of such mechanisms yields taxonomic organization (Kemp & Tenenbaum, 2008; McClelland & Rogers, 2003). In contrast with taxonomic bias accounts, featural learning accounts argue in favor of the gradual emergence of taxonomic organization over the course of development. However, like taxonomic bias accounts, featural learning accounts do not specify any role in semantic organization for spatial or temporal co-occurrences between objects or the words that denote them.

A Potential Role for Statistical Co-Occurrence Regularities

Of the influential accounts reviewed in the previous section, only some restructuring accounts posit any role for co-occurrence regularities in semantic organization. However, even in these accounts, these regularities play a role only early in development, and are subsequently overwritten. At the same time, several findings highlight the potential importance of co-occurrence regularities *throughout* development.

First, evidence from numerous studies suggests that sensitivity to co-occurrence regularities (including the co-occurrence of words and objects) is apparent even early in development (Bulf, Johnson, & Valenza, 2011; Samuelson & Smith, 1999; Wojcik & Saffran, 2015). Moreover, numerous findings attest to effects on children's reasoning of semantic relations that may be formed based on co-occurrence, including *schematic* relations between entities that occur in the same context (e.g., *cow* and *barn*) and *thematic* relations between entities that play complementary roles (e.g., *nail* and *hammer*) (Blaye, Bernard-Peyron, Paour, & Bonthoux, 2006; Fenson, Vella, & Kennedy, 1989; Smiley & Brown, 1979; Tversky, 1985; Walsh, Richardson, & Faulkner, 1993). These findings suggest that accounts of semantic development that do not posit any role for co-occurrence between objects or their labels, such as the taxonomic bias and featural learning accounts, are at best incomplete.

Second, evidence from a handful of studies suggests that semantic relations that may be derived from co-occurrence continue to manifest in semantic organization into adulthood. For example, in a series of ten experiments, Lin and Murphy (2001) observed that relations between entities that adult raters judged as associated in scenes or events (which likely co-occur in the environment) had a pervasive influence on adults' categorization and reasoning that was frequently greater than the influence of taxonomic relations (see also Ross & Murphy, 1999). This evidence is inconsistent with restructuring accounts, in which an influence of co-occurrence early in development is eventually overwritten. More broadly, the proposal inherent in restructuring accounts that later-emerging knowledge and abilities replace those that emerge earlier in development is also inconsistent with evidence that adults revert to childlike patterns of semantic reasoning under cognitive load (Goldberg & Thompson-Schill, 2009).

Finally, the potential contributions of co-occurrence regularities are highlighted by a mechanistic account and corroborating behavioral evidence presented by Sloutsky et al. (2017). This account was inspired by computational modeling evidence that everyday language input, including input to children (Asr, Willits, & Jones, 2016; Frermann & Lapata, 2015; Huebner & Willits, 2018), is rich in statistical co-occurrence regularities that capture links between concepts in semantic organization (see Jones et al., 2015 for a review). First, regularities with which words directly co-occur, such as *fork* and *spaghetti*, link concepts that are reliably associated in semantic knowledge (Hofmann, Biemann, Westbury et al., 2018; Spence & Owens, 1990). Moreover, regularities with which words share each other's patterns of co-occurrence with other words, such as *spaghetti* and *pie* (both co-occur with *fork*), link members of the same taxonomic category. Therefore, according to Sloutsky et al.'s (2017) account, exposure to co-occurrence regularities in language fosters both the learning of associations between concepts whose labels *directly* co-occur, and between taxonomically related concepts whose labels *share* patterns of co-occurrence. However, whereas direct co-occurrence can be gleaned straight from the input and therefore rapidly foster links between concepts, shared patterns of co-occurrence should foster links between members of the same taxonomic category more slowly because they can only be derived by integrating across multiple instances of direct co-occurrence. Moreover, a similar process may unfold for co-occurrence patterns between the entities that concepts represent, given evidence that co-occurrence patterns between words in language are closely mirrored by co-occurrence patterns between objects in everyday visual scenes (Sadeghi, McClelland, & Hoffman, 2015). This account predicts both that: (1) Concepts should be linked based on direct co-occurrence starting early and continuing throughout development, and (2) The linking of taxonomically related concepts should gradually supplement co-occurrence links.

Initial support for this account comes from a series of word learning and lexical extension experiments (Sloutsky et al., 2017) in which both children and adults had to infer a meaning for a novel word embedded in a list of familiar words. When the novel word appeared in a list of words that are associated (and therefore likely to co-occur) with the word "animal", such as "furry" and "zoo", both children and adults inferred that the novel word meant "animal". In contrast, when the novel word appeared in a list of members of the taxonomic category of animals such as "lion" and "bunny", only adults inferred this meaning. Moreover, this account is consistent with and can therefore potentially help explain prior

findings suggesting that taxonomic relations emerge later in development than earlier-emerging relations such as associative links (e.g., Bjorklund & Jacobs, 1985; Blaye et al., 2006; Fenson et al., 1989)

Together, these prior findings suggest that whereas direct co-occurrence-based links may manifest in semantic organization throughout development and into adulthood, taxonomic organization may gradually supplement these links because they are derived (at least in part) from them. However, in addition to being overlooked in traditional theoretical accounts of the development of semantic organization, this possibility has received only limited empirical investigation to date. Furthermore, even when investigated, *actual* co-occurrence in the environment has rarely been assessed. Instead, researchers have investigated semantic relations between items that either are: (1) Judged by the researchers themselves to co-occur, (2) Judged by adult raters to co-occur, or (3) Produced by participants in free association tasks.

Although this approach has yielded evidence that is informative about the development of semantic organization, researchers' or raters' judgments and free associations do not directly estimate co-occurrence in environmental input, because these judgements are themselves *consequences* of learning. This approach therefore does not capture the nature of the input information that contributes to such learning (Hofmann et al., 2018). For example, with respect to researcher and rater judgments, there is evidence that judgments of co-occurrence are contaminated by taxonomic relatedness (and vice versa; Wisniewski & Bassok, 1999). Similarly, the nature of the relations that link words produced in free association tasks must be subjectively inferred and can potentially be taxonomic, derived from co-occurrence, or some other type of link such as a part-whole relation. A more direct estimate of co-occurrence regularities from actual input may therefore provide a more accurate estimate of the role of co-occurrence in semantic organization throughout development.

Current Study

The overall goal of the current study was to investigate the presence of co-occurrence and taxonomic links in lexico-semantic organization across development, from early childhood to adulthood. This investigation was designed to arbitrate between competing theoretical accounts of the development of knowledge organization. Specifically, restructuring accounts predict that co-occurrence should contribute to knowledge organization in childhood, but be replaced by taxonomic relations in adulthood. Neither taxonomic bias nor featural learning accounts make any predictions about the contributions of co-occurrence. However, whereas the former predict that taxonomic relations should contribute from childhood through to adulthood, the latter predict that the contributions of taxonomic relations should substantially increase with age.

A different developmental pattern is predicted by recent proposals that highlight a key role throughout development for co-occurrence in which it both directly fosters relations between concepts, and indirectly fosters relations between concepts that share patterns of co-occurrence and are often taxonomically related (e.g., Sloutsky et al., 2017). Specifically, such proposals predict that relations between concepts whose labels or referents regularly

co-occur should be evident in the semantic organization of both children and adults, and be increasingly supplemented by taxonomic relatedness over the course of development.

We accomplished this investigation by measuring the degree to which familiar concepts were related in young children (4- to 5-year-olds) and adults' semantic knowledge when either the concepts' labels reliably co-occur in linguistic input, or when they are members of the same taxonomic category. To target actual experienced co-occurrence, we identified pairs of items based on the regularity with which the words for a variety of concepts familiar to young children (e.g., *cat*, *table*) co-occurred more reliably with each other than with other words in corpora of child-directed speech. To facilitate the comparison between children and adults, we used paradigms developed to assess semantic relatedness *implicitly*, without the requirement for engaging in and articulating reasoning about relatedness, which adults may accomplish more easily. This approach contrasts with previous studies that have used more explicit reasoning or generalization tasks in an attempt to assess semantic organization across development (e.g., Sloutsky, et al., 2017), in which developmental changes may in part be due to improvements in abilities such as reasoning.

Paradigms.

To obtain a generalizable representation of lexico-semantic organization across development, the three reported experiments used three different paradigms yielding implicit measures of semantic organization. The paradigms all share the same underlying logic: When two concepts are linked, one should automatically activate the other. However, the paradigms measure this automatic activation in different ways.

In Experiment 1, we used a Cued Recall paradigm to measure the effects of co-occurrence and taxonomic relatedness on memory retrieval. The logic of this paradigm was that links between a pair of concepts should facilitate the accuracy with which the words for these concepts are recalled. Co-occurrence and taxonomic links were therefore measured based on the degree to which they facilitated the recall of word pairs, in comparison to unrelated pairs (e.g., Blewitt & Toppino, 1991).

In Experiment 2, we used a Match Verification paradigm in which participants identified whether a word and a subsequent picture denoted the same item (e.g., the word "table" followed by a picture of a table) or different items (e.g., the word "table" followed by a picture of a chair). The logic of this paradigm was that links between pairs of concepts should *interfere* with the ability to say that a word for one concept *does not denote the same item* as a subsequent picture (e.g., Gellatly & Gregg, 1975). Therefore, co-occurrence and taxonomic links were measured based on the degree to which they *interfered* with participants' ability to identify a picture as denoting a different item from its preceding word, relative to unrelated word-picture pairs.

Experiment 3 was designed to provide a more sensitive and graded measure of co-occurrence and taxonomic links. Specifically, we used a variant of Visual World paradigm in which we presented pairs of pictures of familiar, unrelated Target items (e.g., *bed* and *fish*), and measured the degree to which participants looked at a given Target over time upon hearing either a Co-Occur (e.g., *pillow* or *water*), Taxonomic (e.g., *table* or *bird*), or

Unrelated Prime. Unlike the paradigms used in Experiments 1 and 2, this paradigm does not provide a single measure of the relatedness between a given pair of concepts, such as recall accuracy or reaction time. Instead, this paradigm provides a nuanced and graded measure of semantic relatedness: The degree to which hearing the word for one concept influences (over time) looking at a picture of the other concept. As attested by numerous findings, these looking dynamics are sensitive to a variety of relation types, including extremely weak taxonomic relations (Huettig & Altmann, 2005; Huettig, Quinlan, McDonald, & Altmann, 2006; Mirman & Graziano, 2012; Mirman & Magnuson, 2009). Therefore, we measured co-occurrence and taxonomic links based on time course of looking at each Target when it was accompanied by a Co-Occur versus an Unrelated Prime, and a Taxonomic versus an Unrelated Prime.

Experiment 1

Method

Participants.—Informed consent was obtained from parents/guardians of child participants and from adult participants prior to participation. The sample included 31 4–5-year-old children (M_{age}=4.50 years, SD=1.62 years), and 35 Adults (M_{age}=20.16 years, SD=3.66 years). An additional group of seven children and four adults were tested but excluded due to either failure to respond on over a third of trials (six children; three adults), or responding inaccurately on all trials (one child; one adult). An additional eight children completed practice trials only due to failure to reach the accuracy criterion during these trials needed to continue to the experiment (see Procedure below). Children were recruited from families, daycares, and preschools in a metropolitan area of a large Midwestern US city. Adults were undergraduate students at a large Midwestern public university in the same city, and they participated in exchange for partial course credit. These age groups were chosen because: 1) The 4–5 years period is one in which the nature of sematic knowledge remains the subject of active debate, and 2) Comparing semantic organization in early childhood to adulthood affords an investigation into whether early semantic organization is maintained, supplemented, or overwritten by adulthood.

Selection of Candidate Stimuli.—The primary stimuli used in this experiment were word pairs, with each belonging to one semantic Relatedness condition: Co-Occur, Taxonomic, or Unrelated. All pairs were selected (as described below) such that pairs in the Co-Occur condition were words that reliably co-occurred with each other more often than with other words in child speech input, pairs in the Taxonomic condition were words for concepts from the same category with similar meanings according to a database composed by lexicographers of words and their definitions (Wordnet, 2010), and pairs in the Unrelated condition neither reliably co-occurred nor were similar in meaning.

Co-Occurrence Criteria.: The first step taken to select pairs in each condition was to identify a set of words for which lexical norms collected using the MacArthur-Bates Communicative Development Inventory (MB-CDI) were available from WordBank (an open database of children's vocabulary development, Frank, Braginsky, Yurovsky, & Marchman, 2016), and measure their rates of co-occurrence in 25 child speech input corpora from the

CHILDES database (MacWhinney, 2000)[2]. To reduce the computational expense of measuring co-occurrence rates between these words, some classes of words (e.g., all sounds such as "moo") that would *a priori* not be used as stimuli in this research were removed from the full set of words, leaving a list of 538 words. Additionally, to ensure that co-occurrences were measured from child speech *input*, the CHILDES corpora were pre-processed to remove all speech produced by the children themselves. Co-occurrences between these words were then calculated by taking all possible pairs of words within this set, and calculating how frequently they co-occurred with each other within a 7-word window[3] across 25 CHILDES corpora. Finally, to account for the fact that more frequent words co-occur with other words simply by chance, t-scores (Evert, 2008) were calculated for each word pair using the formula below. This formula captures the difference between each word pair's actual, measured co-occurrence frequencies (O), and the frequency of co-occurrence that would be expected by chance given the frequencies of each word in the pair and the size of the combined corpora (E). The larger the difference between observed versus expected frequency, the more reliably words in a pair co-occur:

$$t \, . \, score = \frac{O - E}{\sqrt{O}}$$

Candidate word pairs for use in the Co-Occur condition were then selected as pairs of nouns with t-scores of > 2.5 (following Baayen, Davidson, & Bates, 2008) in which, according to lexical norms accessed from WordBank, both words were produced by at least 80% of 36-month-old children (approximately one year younger than children in our 4-year-old sample).

Taxonomic Criteria.: Taxonomic relatedness was determined based on both the membership of concepts in the same taxonomic category (e.g., clothing, foods, vehicles) and similarity in *meaning* between their labels. To measure similarity in meaning, we measured similarity between the definitions of labels for the items from WordNet (a database of words and their definitions composed by lexicographers). Similarity in WordNet was chosen as the taxonomic relatedness criterion because it captures the essence of taxonomic relatedness – i.e., close similarity in meaning – without relying on judgments of adult participants that may be influenced by non-taxonomic relations (Wisniewski & Bassok, 1999). In WordNet, nouns are first grouped into sets of synonyms, which are in turn linked into a hierarchy according to "IS A" and part-whole relations. Similarity in meaning between pairs of words that label stimuli used in this experiment was measured using the Resnik similarity measure, i.e., the information content (specificity) of the word lowest in the WordNet hierarchy within which the pair of words is subsumed. For example, *dog* and *cat* are subsumed within *carnivore*, whereas *dog* and *kangaroo* are subsumed within mammal; because the information content of carnivore is greater than the information content of mammal (i.e.,

---

[2] The rates of co-occurrence of words in speech input are likely to be similar to the rates with which the objects to which the words refer co-occur (Sadeghi et al., 2015).

[3] This 7-word window was chosen to focus on a span of words that children could plausibly maintain in memory (e.g., Klem, Melby-Lervåg, Hagtvet et al., 2015)

mammal is more abstract), Resnik similarity is higher between dog and cat versus dog and kangaroo.

Candidate word pairs for use in the Taxonomic condition were selected as pairs of nouns with Resnik similarities of >5 in which, as in the selection of candidate word pairs in other conditions, both words were produced by at least 80% of 36-month-old children according to WordBank production norms. The rationale of the Resnik similarity criterion of >5 is illustrated in Figure 1, which shows that the similarity between Taxonomic pairs and other items from the same taxonomic category (e.g., clothing) are above 5, whereas the similarity between items from different taxonomic categories are substantially below this value.

Unrelated Criteria.: Candidate Unrelated word pairs were pairs of nouns that met the same WordBank production norm criterion as candidates in the Co-Occur and Taxonomic conditions, with t-scores of < 1.5 and Resnik similarities of < 4.

Composition of Stimulus Set.—From the sets of candidate pairs, eight pairs were selected for each of the three Relatedness conditions (Co-Occur, Taxonomic, and Unrelated, for a total of 24 pairs) such that: (1) Pairs in the Co-Occur condition did not meet the Taxonomic Criteria and pairs in the Taxonomic condition did not meet the Taxonomic Criteria, (2) The mean percentage of 36-month-olds who produced the words in the pairs according to Wordbank norms was equated across conditions, and (3) No words appeared in more than one condition (see Table 1 for all 24 pairs, and Appendix A for t-score and Resnik similarity measures for each pair). Four additional nouns that met the WordBank production norm criterion were selected to construct pairs used for demonstration and practice (see Procedure below). All words were recorded by both a male and a female speaker using an engaging, child-friendly intonation.

Design.—The relatedness condition was manipulated within subjects, with each pair presented in Table 1 occurring only in one condition. Because pilot testing indicated that 12 pairs was the maximum number that could be presented to children without producing floor effects, the total set of 24 pairs (i.e., eight pairs in each of the three Relatedness conditions) was divided into two Stimulus Sets. Accordingly, each Stimulus Set contained 12 pairs, with four pairs in each of the three Relatedness conditions. In all word pairs, each word in a pair was randomly assigned to be either the Cue or to-be-remembered Target. Across conditions, Cue words were presented using the male speaker's voice, and Targets using the female's voice. Additionally, the 12 word pairs in a Stimulus Set were pseudorandomized into three blocks, such that: (1) Each pair only appeared once in the entire Stimulus Set, and (2) Each block contained 1–2 pairs from each condition. The order of these blocks was counterbalanced across participants.

Procedure.—In all experiments reported here, participants were presented with procedures approved by The Ohio State University Institutional Review Board (Protocol #: 2004B042, *Comprehensive protocol for cognitive development research*). Adults were tested in a quiet space in the lab on campus, and children were tested either in a quiet space in the lab, or at their preschool or daycare. The procedure was similar for adults and children, with the following exceptions: 1) The instructions were conveyed verbally by a hypothesis-blind

experimenter for children, and as text on a computer screen for adults, and 2) Children made verbal responses recorded by the experimenter, whereas adults typed their responses.

To start, participants were introduced to two sock puppets depicted on the computer screen, named Izzy and Ozzy. Participants were informed that they were going to play a game with Izzy and Ozzy, in which Izzy and Ozzy would say pairs of words (children were given an additional explanation about what a "pair" is). The two unrelated Cue-Target word pairs selected for demonstration/practice were then played sequentially, while animations depicted one puppet saying the Cue word, and the other saying the Target word. Next, participants were told that they were going to listen to the word pairs again, but to pay close attention to the pairs of words that go together, because it would then be their job to pretend to be Ozzy and either say (children) or type (adults) the word that went with Izzy's word. Participants then proceeded to complete practice rounds with the same two unrelated Cue-Target word pairs. Each practice round consisted of: 1) A "Study" phase, in which the two word pairs were presented as spoken by Izzy and Ozzy, and 2) A "Test" phase, in which only the Cue in each pair was presented as spoken by Izzy, and participants were prompted to either say or type the Target that had been spoken by Ozzy (Figure 2). Participants received feedback about whether their responses were correct or incorrect. Participants completed up to three practice rounds until they either responded with the correct Target for both Cues within around, or the experiment was terminated due to failure to reach this criterion.

Participants then proceeded to complete the three blocks of Cue-Target pairs in the Stimulus Set to which they had been randomly assigned. Each block followed the same Study and Test phase format as the practice rounds, with the exception that participants did not receive feedback about the accuracy of their responses. At the beginning of the Test phase of each block, participants were encouraged to take their best guess when they were unsure of the correct answer. The full experiment took approximately 7–10 minutes for adults, and 10–12 minutes for children.

Results and Discussion

The primary outcome measure of interest for this study was the accuracy with which participants recalled Target words paired with Cues in each of the three Relatedness conditions: Co-Occur, Taxonomic, and Unrelated. Responses were scored as accurate when participants made responses identical to the Target or morphological variants of the Target (e.g., "spoons" instead of "spoon"). Additionally, three responses (all in children) in which the correct Target was "street" and the child responded "road" were also scored as accurate (all reported analyses produce the same results when these responses are excluded). No other cases of responses synonymous with the Target occurred.

All analyses were conducted in the R environment. Mixed effects models and corresponding χ^2 or F-statistics were generated using the lme4 (Bates, Maechler, Bolker, & Walker, 2015) and car (Fox & Weisberg, 2011) packages, respectively.

Preliminary Analyses: Stimulus Set Comparison.—Prior to comparing accuracy in the Semantic Relatedness conditions, we first tested whether any effect of condition varied across the two Stimulus Sets in children and adults. Specifically, for data from each age

group, we generated a binomial generalized linear mixed effects model with Accuracy as the outcome variable, Relatedness condition (Co-Occur, Taxonomic, and Unrelated) and Stimulus Set (1 versus 2) as fixed effects, and random intercepts for participant and item. This analysis revealed no significant interaction between Relatedness condition and Stimulus Set ($ps > .09$). For all subsequent analyses, we therefore collapsed across Stimulus Sets.

Primary Analyses.—Memory accuracies by age and condition are presented in Figure 3. To test the relative influences of Semantic Relatedness conditions on recall accuracy, we generated an omnibus binomial generalized linear mixed effects model with Accuracy as the outcome variable, Relatedness condition (Co-Occur, Taxonomic, and Unrelated) and Age group (children and adults) as fixed effects, and random intercepts for participant and item. This analysis revealed main effects of both Relatedness condition ($\chi^2(2)=20.35$, $p<.001$) and Age group ($\chi^2(1)=15.74$, $p<.001$) that were qualified by an interaction between them ($\chi^2(2)=9.38$, $p=.009$).

To investigate the interaction between Relatedness condition and Age group, we compared the effects of the different Relatedness conditions in each Age group.

Relatedness conditions in Each Age Group.: To compare the effects of the Relatedness conditions in each Age Group, for each Age group, we first generated a binomial generalized linear mixed effects model with Accuracy as the outcome variable, Relatedness condition as a fixed effect, and random intercepts for participant and item. These models revealed significant effects of Relatedness condition in each age group ($ps < .01$) (Figure 3). To conduct pairwise comparisons of the Relatedness conditions in each age group, we re-generated the model for each age with each of the Relatedness conditions as the reference level, and applied Bonferroni-adjustments to the resulting p-values.

In children, these analyses revealed significant differences between the Co-Occur and both Unrelated and Taxonomic conditions ($ps < .001$), but no difference between the Taxonomic and Unrelated conditions ($p > .99$). In adults, these analyses revealed a significant difference between the Co-Occur and Unrelated conditions ($p = .003$), the Taxonomic and Unrelated conditions ($p=.04$), and no significant difference between Co-Occur and Taxonomic conditions ($p>.99$).

Individual Differences.—The results of the primary analyses suggest that Co-Occurrence links manifested in semantic organization in both children and adults, whereas Taxonomic relatedness manifested only in adults' semantic organization. However, it is important to highlight that the lack of a taxonomic influence in young children observed in the present experiments is a null finding, from which strong conclusions cannot be drawn. For example, an influence of taxonomic relatedness may have been present, but was too small in magnitude and/or transpired in too few children to detect. We therefore investigated this possibility using a qualitative analysis of the magnitudes of Co-Occur and Taxonomic relatedness effects within individuals in both the child and adult samples.

In these analyses, we quantified the magnitude of both Co-occurrence and Taxonomic effects for each participant by calculating both a Co-Occur and a Taxonomic Difference Score based on the difference between each of these conditions and the Unrelated condition (such that a Difference Score of 0 for a given condition indicates no influence of the condition on behavior). The densities of the distributions of Difference Scores in each age group in each experiment are depicted in Figure 4. As in the primary analyses, these distributions show an influence of Co-occurrence relatedness in children, and both Co-occurrence and Taxonomic relatedness in adults. However, these distributions also suggest an influence of Taxonomic relatedness was present in children, but both occurred in fewer participants and tended to be smaller in magnitude than the influence of Co-Occurrence relatedness.

Summary.—The results of Experiment 1 revealed a substantial influence of co-occurrence regularities in both young children and adults, such that co-occurrence between to-be-remembered Cue and Target word pairs facilitated subsequent recall. In contrast, taxonomic relatedness did not significantly affect recall in children. Instead, the influence of taxonomic relatedness transpired only in adults. However, our quantification of co-occurrence and taxonomic effects within individuals adds nuance to this pattern. Specifically, this qualitative analysis both corroborates these results, and suggests that taxonomic contributions, rather than being totally absent in children, were instead present but too weak and uncommon to produce significant effects at the group level.

These results highlight the role of co-occurrence in semantic organization throughout development. Moreover, these results also suggest that over development, new (and perhaps more advanced) taxonomic organization *increasingly supplements* co-occurrence rather than replaces it (see also Supplementary Materials for quantifications of positive correlations between co-occurrence and taxonomic effects, consistent with the proposal that taxonomic relations build upon co-occurrence). To examine the generalizability of this finding, we conducted Experiment 2.

In Experiment 2, we used an entirely different paradigm to investigate participants' sensitivity to co-occurrence and taxonomic relatedness. In this paradigm, participants were presented with word-picture pairs in which the picture either did or did not depict the item referred to by the word (e.g., the word "lion" followed by a picture of a lion or the word "bottle" followed by a picture of a baby). In contrast to examining whether relatedness *improved* performance (as was done in Experiment 1), we used this paradigm to measure the degree to which co-occurrence or taxonomic relatedness between word-picture pairs *interfered* with participants' ability to indicate that the picture did *not* depict the item labeled by the word.

Experiment 2

Method

Participants.—Informed consent was obtained from parents/guardians of child participants and from adult participants prior to participation. The sample included 41 4–5-year-olds (M_{age}=4.05 years, SD=1.71 years) and 42 Adults. Two additional children were

tested but excluded due to mean reaction times more than two standard deviations above the mean reaction time for this age group. An additional three children completed practice trials only due to failure to reach the accuracy criterion needed to continue to the experiment (see Procedure below). Children were recruited from families, daycares, and preschools in a metropolitan area in a Midwestern US city. Adults were recruited from the undergraduate population at a public university in the same city and participated in exchange for partial course credit.

Stimuli and Design.—The primary stimuli were similar to those used in Experiment 1, with the following changes. First, we added several pairs to those used in Experiment 1. Specifically, we added one pair to the Co-Occur, Taxonomic, and Unrelated conditions, for a total of 9 pairs in each condition. Additionally, from the list of nouns not used in the Co-Occur, Taxonomic, or Unrelated conditions, an additional 24 nouns that met the production by 80% of 36-month-olds WordBank criterion were selected for use in an Identical condition in which a "pair" consisted of a word and a picture depicting the same thing (e.g., the word "lion" followed by a picture of a lion). The Co-Occur, Taxonomic, Unrelated, and Identical pairs each appeared once in the experiment, for a total of 51 pairs. These pairs were pseudo-randomized prior to the experiment such that no more than two pairs from the same condition appeared consecutively. An additional 18 nouns that met the WordBank production norm criterion were also selected to appear as demonstration and practice stimuli (see Procedure below).

Second, whereas the stimuli were divided into separate Stimulus Sets in Experiment 1, all stimuli were presented to all participants in this experiment. Finally, to eliminate potential effects of perceptual similarity between members of the same pair, both spoken words and pictures were generated for all words, such that a spoken word was used for one member of the pair and a picture for the other member. Specifically, for each pair, one word was randomly assigned to appear in the experiment as a spoken word, and the other was assigned to appear as a picture. The spoken word version was recorded by a male speaker using an engaging, child-friendly intonation, and the picture version was a color photograph of the item isolated on a white background (resized to 276×276pi).

As in Experiment 1, the relatedness condition varied within subjects. In addition, also as in Experiment 1, each word pair was presented only once.

Procedure.—Adults were tested in a quiet space in the lab on campus, and children were tested either in a quiet space in the lab, or at their preschool or daycare. The procedure was similar for adults and children, with the exceptions that: 1) The instructions were conveyed verbally by a hypothesis-blind experimenter for children, and as text on a computer screen for adults, and 2) Children chose response options using a touchscreen, whereas adults used a mouse.

To start, participants were instructed that they were going to play a game in which they would hear a word and then see a picture, and that their job was to identify whether the picture was "of the same thing" as the word. Participants were then instructed to click (adults) or touch (children) a smiley face depicted on the bottom of the screen if the picture

was of the same thing as the word. Two Demonstration trials were then presented. In these trials, a word was followed by a picture of the same thing (e.g., the word "pretzel" followed by a picture of a pretzel), and the smiley face was highlighted as the correct response. Subsequently, participants were instructed to click or touch a frowny face also depicted at the bottom of the screen if the picture was *not* of the same thing as the word, and then shown Demonstration trials using two unrelated word-picture pairs (e.g., the word "zebra" followed by a picture of scissors). Participants then proceeded to complete eight Practice trials composed of an equal number of word-picture pairs in which the picture was of the same thing as the word, and pairs in which the picture was of an item unrelated to the word. Participants were encouraged to respond as quickly and accurately as possible. The smiley and frowny face response options only appeared 250ms after the onset of the picture, and remained on the screen for 6 seconds, to impose a time limit on the window within which responses could be made. Participants received corrective feedback after each trial telling them whether they were correct, incorrect, or too slow (if they failed to respond during the time limit). If a participant failed to reach a criterion of 5 out of 8 trials correct, the Practice trials repeated up to two additional times until the criterion was reached. If a participant failed to reach the criterion after three rounds of Practice trials, the experiment was terminated for that participant.

Participants then proceeded to complete the experimental trials, in which items from the Co-Occur, Taxonomic, Unrelated, and Identical conditions were presented in a pseudo-randomized order, such that no more than two pairs from the same condition were presented consecutively. These trials followed the same format as Practice trials, with the exceptions that the response options remained on the screen until participants made a response (i.e., no upper time limit was imposed), and no feedback was provided. The full experiment took approximately 4–5 minutes for adults, and 5–8 minutes for children.

Results and Discussion

In the experiment, accurate responses were those in which participants responded that the word and picture were of the same thing (henceforth referred to as "yes" responses) in the Identical condition, or that the word and picture were *not* of the same thing (henceforth referred to as "no" responses) in all other conditions. Prior to conducting hypothesis-testing analyses, we first determined that both children and adults understood the task: In both age groups, overall response accuracies for both "yes" and "no" responses were significantly above chance (Children: M_{yes} = 83.94%, M_{no} = 88.41%; Adults: M_{yes} = 97.62%, M_{no} = 98.54%; all ps < .001).

Primary Analyses.—Reaction times by age and condition are presented in Figure 4. Our primary measure of interest was how much more difficult it was for participants to make accurate "no" responses to non-identical word-picture pairs in the Co-Occur and Taxonomic conditions compared to the Unrelated condition. We measured comparative difficulty using reaction time (log-transformed for analyses) in the three conditions. Such interference effects in the Co-Occur or Taxonomic conditions were taken as evidence that a participant was sensitive to the respective relation.

To test the relative influences of the Relatedness conditions on reaction time, we generated an omnibus linear mixed effects model with Reaction Time as the outcome variable, Relatedness condition (Co-Occur, Taxonomic, and Unrelated) and Age group (children and adults) as fixed effects, and a random intercept for participant. This analysis revealed main effects of both Relatedness condition, $F(2,1977.26)=21.89$, $p<.001$ and Age, $F(1, 81.01)=248.66$, $p<.001$, which were qualified by an interaction, $F(2,1977.51)= 11.29$, $p<.001$.

Relatedness conditions in Each Age Group.: To investigate the interaction between Relatedness condition and Age group, for each Age group, we first generated a linear mixed effects model with Reaction Time as the outcome variable, Relatedness condition as a fixed effect, and participants as a random effect. These models revealed significant effects of Relatedness condition in each age group ($ps < .05$).

To conduct pairwise comparisons of the Relatedness conditions in each age group, as in analyses for Experiment 1, we re-generated the model for each age group with each of the Relatedness conditions as the reference level, and applied Bonferroni-adjustments to the resulting p-values. In children, these analyses revealed significant differences between the Co-Occur and both Unrelated and Taxonomic conditions ($ps < .001$), but no difference between the Taxonomic and Unrelated conditions ($p > .99$). In adults, these analyses revealed a significant difference between the Co-Occur and Unrelated conditions ($p = .014$), no significant difference between Co-Occur and Taxonomic conditions ($p=.352$), and no significant difference between the Taxonomic and Unrelated conditions ($p=.630$). However, there was a numerical trend for longer reaction times in both Co-Occur versus the Taxonomic condition, and for the Taxonomic versus the Unrelated Condition ($M_{\text{co-occur}} = 974ms$, $M_{\text{taxonomic}} = 945ms$, $M_{\text{unrelated}} = 928ms$).

Individual Differences.—We supplemented our primary analyses by following the same approach as in Experiment 1 to quantifying the magnitudes of Co-Occur and Taxonomic relatedness effects within individuals in both the child and adult samples. As shown in Figure 6, as in Experiment 1, this analysis both corroborates the results of our primary analyses, and suggests that taxonomic relatedness effects were present but too weak and uncommon in children to reach significance at the group level.

Summary.—The results of this experiment provided further evidence for a substantial sensitivity to co-occurrence regularities that manifested in young children, and continued into adulthood. Specifically, in both age groups, participants found it more difficult to identify when a picture did *not* depict the same thing as a preceding word if the word and the picture's label reliably co-occur in linguistic input.

In contrast, sensitivity to taxonomic relatedness in this task did not reach significance in young children. As in Experiment 1, our qualitative analysis of individual co-occurrence and taxonomic effects suggests that the absence of the taxonomic effects at the group level in children was due to the weakness and rarity of these effects in children, rather than to their complete absence. In adults, although the influence of taxonomic relatedness was not significantly smaller than the influence of co-occurrence, responses to taxonomic pairs also

did not significantly differ from responses to unrelated pairs. This replication of the contribution of co-occurrence in children and adults using two very different paradigms underscores the significance of sensitivity to co-occurrence regularities in relational knowledge across development (see also Supplementary Materials for quantifications of positive correlations between co-occurrence and taxonomic effects, consistent with the proposal that taxonomic relations build upon co-occurrence).

The purpose of Experiment 3 was to both test the generalizability of these patterns to another very different paradigm, and to gain a more sensitive and nuanced measure of co-occurrence and taxonomic links. Specifically, our qualitative analyses for both Experiments 1 and 2 suggested taxonomic links in young children that were too weak to reach significance at the group level. Therefore, in Experiment 3, we used a paradigm that has been shown to yield a sensitive and graded measure of even weak semantic relations: The Visual World paradigm.

In the Visual World paradigm, participants view items (typically pictures) while hearing linguistic input, such as a word. Numerous studies have provided evidence that individuals tend to look at pictures that are semantically related to the words that they hear (e.g., Huettig & Altmann, 2005; Mirman & Magnuson, 2009). Accordingly, the degree to which hearing a word for one concept (e.g. *cat*) prompts looking at a picture of a semantically related concept (e.g., *dog*) can serve as a measure of the degree to which the concepts are linked in an individual's semantic knowledge. A similar tendency has been observed in infants in preferential looking paradigms (Arias-Trejo & Plunkett, 2009; Bergelson & Aslin, 2017) suggesting that measures of this looking behavior are appropriate for a wide developmental age range.

Critically, unlike the paradigms used in Experiments 1 and 2, this paradigm does not yield only a single snapshot measurement of relatedness between two concepts. Instead, it measures the *degree* to which one concept (presented as a word), activates another concept (presented as a picture). Moreover, it measures how this degree of activation unfolds *over time* in the milliseconds following the presentation of the word. This paradigm therefore yields a graded, nuanced measure that has been shown to be sensitive to even weak taxonomic relations (Mirman & Magnuson, 2009).

Experiment 3

Overview

For Experiment 3, we developed a variant of the Visual World paradigm with key characteristics designed to probe the degree to which words activate co-occurring and taxonomically related concepts directly. In this paradigm, participants heard "Prime" words while freely visually inspecting visual displays containing two "Target" pictures of unrelated familiar items (e.g., *bed* and *fish*). A given pair of Target pictures always appeared with each other, and never with other items. Across presentations of a given Target pair, we varied whether the Prime word was: (1) A Co-Occur Prime that co-occurred with one of the Targets (e.g., *pillow* or *water*), (2) A Taxonomic Prime that was taxonomically related with one of the Targets (e.g., *table* or *bird*), or (3) An Unrelated Prime that was unrelated to both Targets.

Following presentation of the Prime, participants freely viewed the Targets for 2000ms. We measured the activation of a Target by a Prime based on the degree to which participants looked more at a given Target *over time* when accompanied by a Co-Occur or Taxonomic versus an Unrelated Prime.

It is worth highlighting two characteristics that distinguish this version of the Visual World paradigm from the ways in which this paradigm has typically been implemented in prior research. First, whereas prior approaches have manipulated semantic relatedness by manipulating the *pictures* that appear with a given Prime word (e.g., presenting "bed" with either a picture of a pillow or a chair), our version manipulated the Prime word that was presented with a given pair of pictures. This approach allowed us to measure the temporal dynamics with which the concepts depicted by the Target pictures were activated upon hearing different Primes while keeping the pictures themselves constant, and therefore avoiding contamination from visual salience, visual interest, and so on. Second, in trials in which participants heard a Co-Occur, Taxonomic, or Unrelated Prime, participants did not complete a task, and instead freely viewed the Targets. This characteristic kept our measure of semantic relatedness between Primes and Targets *implicit*, as in Experiment 1 and 2. Instead, these trials were interspersed with trials of a cover task, in which participants heard the word "yellow" or "blue", and clicked a button on a button box of the corresponding color.

Method

Participants.—Informed consent was obtained from parents/guardians of child participants and from adult participants prior to participation. The sample included 36 4–5-year-olds (M_{age}=4.43 years, SD=0.32 years) and 37 Adults. Children were recruited from families, daycares, and preschools in a metropolitan area in a Midwestern US city. Adults were recruited from the undergraduate population at a public university in the same city and participated in exchange for partial course credit.

Stimuli.—The primary stimuli were similar to those used in Experiments 1 and 2. The primary difference in this experiment was that, instead of separate sets of Co-Occur, Taxonomic, and Unrelated pairs, we constructed sets in which a Target was combined with both a Co-Occur Prime and a Taxonomic Prime. Further, we organized these sets into pairs ("Pair Sets") in which: (1) The Targets in the Pair Set were both unrelated and approximately equivalently familiar (i.e., were produced by a similar percentage of 36-month-old children according to production norms), and (2) The Primes for one Target in a Pair Set were unrelated to the other Target (Table 2 and Appendix B). Each Pair Set additionally included one Unrelated Prime that was unrelated to both Targets.

Primes were presented as words recorded in the same manner as in Experiments 1 and 2. We additionally recorded the words "yellow" and "blue" for use in cover task filler trials (see Procedure). Targets were presented as pictures each subtending approximately 5.3° of visual angle.

Finally, the total number of items was expanded in this experiment following the same co-occurrence and taxonomic criteria as in Experiments 1 and 2, for a total of 22 sets organized into 11 Pair Sets[4]

Design.—As in Experiments 1 and 2, the Relatedness condition (Co-Occur, Taxonomic, and Unrelated) varied within subjects, and Age varied between subjects. Within a block of trials, there were a total of 88 trials comprised of 22 trials of each of the following four types: (1) Co-Occur (each of the 11 Target pairs were presented with the two Co-Occur Primes from their Pair Set), (2) Taxonomic (each of the 11 Target pairs were presented with the two Taxonomic Primes from their Pair Set), (3) Unrelated (each of the 11 Target pairs were presented twice with the Unrelated Prime from their Pair Set), and (4) Filler (each of the 11 Target pairs were presented with the words "yellow" and "blue"). Children completed a single block of trials, and adults completed two blocks.

Apparatus.—This experiment used an EyeLink Portable Duo eye tracking system that measures eye gaze by computing the pupil-corneal reflection at a sampling rate of 500Hz. We additionally constructed a non-functional "button box" with yellow and blue buttons for use in the cover task that participants completed during the experiment (see Procedure).

Procedure.—Adults were tested in a quiet space in the lab on campus, and children were tested either in a quiet space in the lab, or at their preschool or daycare. The procedure was similar for adults and children, with the exception that children completed one block of trials, and adults completed two blocks (i.e., repeated the same block twice with randomized trial orders).

Following calibration of the eye tracker, participants began a practice session of a cover task. The purpose of this cover task was to keep participants engaged in looking at the screen and listening to the words, but consisted only of filler trials that were not analyzed. Specifically, to perform the cover task, participants were given the non-functional button box and told that they would see two pictures: One on a yellow background on the left, and one on a blue background on the right. They would then hear either the word "yellow", or the word "blue", such that their job was to click the yellow button on the button box if they heard "yellow", and the blue button if they heard "blue". The practice session consisted of 11 trials of this cover task. In each trial, the two pictures on the screen were the two Targets from one of the 11 Pair Sets. In cover task filler trials only, the experimenter terminated the trial upon observing the participant clicking one of the buttons. The timing of events in cover task and subsequent experimental trials is depicted in Figure 7.

Following completion of the practice session, participants were informed that they would continue to play the same game, but that it would get "a bit tricky", because sometimes they would hear a word that was neither blue nor yellow. Participants were instructed to not click either of the buttons if this occurred.

---

[4]To generate this larger stimulus set, we relaxed the familiarity criterion such that words only needed to be produced by at least 55% of 36-month-olds. However: (1) The majority (86%) of words still met the 80% criterion used in Experiments 1 and 2, and (2) The average production norm value was equated across Co-Occur, Taxonomic, and Unrelated Primes (all produced by ~89% of 36-month-olds).

Participants then proceeded to complete either one block of trials (children) or two blocks (adults). On each trial, the two pictures on the screen were the two Targets from one of the Pair Sets. To create the three Relatedness conditions, in experimental trials, the word was either: (1) The Co-occurrence Prime for one of the Targets, (2) the Taxonomic Prime for one of the Targets, or (3) The Unrelated Prime for the Targets.

These experimental trials were randomly ordered and interspersed with the above described cover task filler trials, in which the word was either "yellow" or "blue". The pairs of Targets in the Pair Set were each presented twice in filler trials, once with the word "yellow" and once with "blue". In combination with the experimental trials, the Targets from each Pair Set were therefore presented a total of 8 times within a block (twice in each of the three Relatedness conditions and twice in filler trials), within which the locations of the Targets were counterbalanced. The full experiment took approximately 10–12 minutes for children and adults.

Using this procedure, we measured the degree to which the looking dynamics for the two Target pictures in a Pair Set varied according to the relation between each Target and the Primes. The fact that each pair of Target pictures was always presented together across the Prime relatedness conditions allowed us to control for effects of visual features, salience, similarity, etc. while measuring these looking dynamics.

Results and Discussion

To test the contributions of co-occurrence and taxonomic relatedness in children and adults, the data from this experiment were used to compare the time course of looking at Targets accompanied by Co-Occurring or Taxonomic Primes versus Unrelated Primes in children and adults. To conduct this comparison, we first processed the raw eye tracking data to generate outcome variables of interest.

Outcome Variables.—Data from practice and filler trials were excluded from analyses. The raw eye tracking data consisted of the position of gaze on the screen sampled every 2ms within experimental trials, which was identified as falling within an AOI for the image on the left, an AOI for the image on the right, or within neither AOI. We first removed data from the 500ms prior to the onset of the word, then divided the remaining two seconds into 100ms time bins. We used these data to generate two outcome variables.

<u>**Target Dwell Time.:**</u> We first calculated a "Target Dwell Time" value for each Target in each time bin in the Co-Occurring, Taxonomic, and Unrelated Prime conditions. This Target Dwell Time value captured the *amount of time spent looking at the Target in each time bin when it was accompanied by a Co-Occurring, Taxonomic, or an Unrelated Prime*. These values were used to test whether the time course of looking at a Target differed when accompanied by a Co-Occur or Taxonomic versus an Unrelated Prime in children and adults (for analyses of the *proportion* of dwell time for each Target, out of the total dwell time to both Targets in a Pair Set, see Supplemental Materials).

<u>**Difference from Unrelated.:**</u> To test for differences in the degree to which the Co-Occur versus Taxonomic Prime conditions *deviated from the Unrelated Prime condition*, we

calculated a "Difference from Unrelated" value for each Target in each time bin. We calculated this value by subtracting the Unrelated Target Dwell Time for a Target/time bin from both the corresponding Target Dwell Time in the Co-Occur condition, and the Target Dwell Time in the Taxonomic condition. The Difference from Unrelated value therefore captures *the degree to which participants looked more at each Target in each time bin when it was accompanied by a Co-Occur or a Taxonomic Prime than when it was accompanied by an Unrelated Prime* (for comparable analyses of *proportion* of Target looking in each time bin, see Supplemental Materials).

Analysis Approach.—We followed the Growth Curve Analysis (GCA) approach developed by Mirman and colleagues (Mirman, Dixon, & Magnuson, 2008) to analyze our data. The GCA approach involves the generation of hierarchical mixed effects models, starting with a "base" model that captures how looking behavior changes over time overall, without considering variation across individuals or experimental manipulations. First, the intercept that captures the average value of the outcome variable. In addition, the base model also includes a linear term that captures monotonic changes in the value of the outcome variable over time, and a quadratic term that captures the sharpness of the peak in looking over time. Finally, cubic and quartic terms capture changes in the asymptotic tails of the outcome variable change over time that are not typically informative about the influences of experimental manipulations (Mirman et al., 2008).

To analyze the effects of experimental manipulations, the base model is supplemented with: Fixed effects of experimental manipulations, random intercepts for participants (and items if appropriate), and random slopes for effects of experimental manipulations within participants (and items, if appropriate). The interpretation of significant fixed effects on the model terms are as follows: (1) Effects on the intercept capture overall effects collapsed across the entire time period on the outcome variable; (2) Effects on the linear term capture effects on the rate of linear change in the outcome variable, similar to linear regression; and (3) Effects on the quadratic term capture effects on the sharpness of the peak with which the outcome variable increases and then decreases (or decreases then increases).

Target Dwell Time Analysis.—We first tested whether the temporal dynamics of looking at Targets differed when accompanied by Co-Occur or Taxonomic Primes in comparison to when accompanied by Unrelated Primes. Specifically, we generated separate models of Dwell Times for Targets in each time bin for children and adults that both supplemented the base model with a fixed effect of Relatedness condition (with Unrelated as the reference level to which Co-Occur and Taxonomic were compared). These models additionally included random intercepts for participant and item, and random slopes for the effect of Relatedness condition within participants and within items.

The parameter estimates and their significance levels are reported in Table 3. Both children and adults looked more overall at a given Target when they heard either a Co-Occur or a Taxonomic versus an Unrelated Prime (as shown by significant effects on the Intercept). Co-Occur and Taxonomic Primes also affected changes in looking at a given Target over time, including the rate at which looking at the Target increased (Linear term) and/or the sharpness of the peak in Target looking time (Quadratic term). Taken together, these results

show that concepts depicted by Targets were activated by both Co-Occur and Taxonomic Primes in both adults and children (see Supplemental Materials for similar results from analyses of Target dwell proportions).

However, this analysis does not reveal the *relative* contributions of Co-Occur versus Taxonomic Primes. To *compare* the contributions of co-occurrence and taxonomic relatedness, in the following analysis, we compared Difference from Unrelated in the Co-Occur and Taxonomic conditions.

Difference from Unrelated Analysis.—This analysis tested whether there was a difference in the degrees to which the Co-Occur versus Taxonomic conditions deviated from the Unrelated condition in children and adults. Specifically, we generated separate models of Difference from Unrelated values for children and adults that both supplemented the base model with a fixed effect of Relatedness condition (Co-Occur and Taxonomic only), random intercepts for participant and item, and random slopes for the effect of Relatedness condition within participants and within items. Figure 8 depicts the Difference from Unrelated data and the corresponding fitted data from the models.

The parameter estimates and their significance levels are reported in Table 4. In children, Co-Occur Primes produced grater rates of increased looking at Targets (relative to Unrelated Primes) than Taxonomic Primes. In contrast, in adults, no such differences were observed: Co-Occur and Taxonomic Primes affected looking at Targets relative to Unrelated Primes to equivalent extents (for similar results from analyses of Target proportions, see Supplemental Materials).

Summary.—The results of this experiment provided further nuance to our picture of the developmental trajectory of semantic organization. First, this experiment revealed that an influence of co-occurrence that persisted from early childhood to adulthood, corroborating results from Experiments 1 and 2.

Critically, this experiment revealed an influence of taxonomic relatedness that was initially weaker than co-occurrence in young children, but became similar in magnitude to co-occurrence by adulthood. This result explicitly captures and quantifies the developmental trajectory suggested by our analyses of individual differences in Experiments 1 and 2, in which taxonomic relatedness in the course of development supplements co-occurrence-based links.

General Discussion

Across experiments that used three different paradigms to yield implicit measures of semantic organization, we observed substantial effects of co-occurrence in both young children and adults. In contrast, the data suggest that taxonomic relatedness increasingly supplements co-occurrence with development. Importantly, due to the implicit nature of the measures of semantic organization used in these experiments, this developmental pattern is unlikely to be attributable to other developmental changes, such as improvement in explicit reasoning abilities.

These findings arbitrate between the predictions of different accounts of semantic organization development. First, the evidence for a continued contribution of co-occurrence to semantic organization in adults is inconsistent with Restructuring accounts, which predict that early-emerging organization based on environmental input (such as co-occurrence) is later overwritten by taxonomic relations. Second, the substantial contributions of co-occurrence to semantic organization throughout development suggest that accounts that do not posit any role for co-occurrence, including both Taxonomic Bias and Featural Learning accounts, are at best incomplete. Specifically, although the sources of input to semantic organization highlighted by these accounts – e.g., labels in Taxonomic Bias accounts and features in Featural Learning accounts – may indeed contribute to semantic organization, co-occurrence regularities also appear to play a key role that these accounts overlook.

The present findings are instead most consistent with a recent mechanistic account proposed by Sloutsky et al. (2017). According to this account, co-occurrence contributes to semantic organization from early in development onward because it is directly observable from environmental input. Taxonomic relations then increasingly come to contribute to semantic organization as they are *derived from* regularities with which different labels share patterns of co-occurrence with each other (e.g., members of the same taxonomic category such as *spaghetti* and *pie* share each other's patterns of co-occurrence with *fork*, *plate*, etc.). The developmental trajectory predicted by this account, in which co-occurrence contributes to semantic organization throughout development and is gradually supplemented by taxonomic relations, was corroborated by the results of the present experiment.

In principle, other, as-of-yet unproposed accounts could also explain the present findings as the result of two entirely separate processes for forming co-occurrence-based and taxonomic relations that develop asynchronously. For example, the more gradual emergence of taxonomic relations might be interpreted as resulting from a gradually-emerging sensitivity to the features that members of taxonomic categories share (e.g., Sloutsky, 2010; Smith & Heise, 1992). Alternately, the gradual emergence of taxonomic relations might be driven by learning both words such as "animal", "clothes", "furniture" etc., and to infer that these denote stable taxonomic categories (e.g., Fulkerson & Waxman, 2007; Gelman & Coley, 1990; Gelman & Markman, 1986). However, regardless of the theoretical framework within which they are interpreted, the findings nonetheless underline the importance of incorporating a key role for co-occurrence regularities in any account of semantic development.

However, the trajectory of semantic organization development cannot be inferred from the present experiments alone. To contextualize these findings, we next evaluate the degree to which this developmental trajectory is consistent with evidence from prior research on semantic development. In this evaluation, we highlight how the present findings are both consistent with, and expand upon much of the large body of prior semantic development research. Finally, we discuss potential mechanistic explanations for the developmental trajectory observed in the present experiments that represent targets for future research.

Developmental Trajectories Observed in Present and Prior Research

Contribution of Co-Occurrence.—Across the three experiments, we observed significant contributions of co-occurrence to semantic organization from early childhood into adulthood. In both young children and adults, co-occurrence: (1) Improved recall of word pairs, (2) Interfered with the ability to identify a picture as not of the same thing as a preceding word, and (3) Guided the dynamics of looking behavior. In this section, we evaluate these findings in the context of prior research. Although a role for co-occurrence throughout semantic organization development has been overlooked or posited to be transient in the majority existing accounts, the evidence supporting this role provided in the present experiments is consistent with many prior findings.

First, our evidence that co-occurrence contributes to semantic organization throughout development is consistent with numerous findings from statistical learning research. Specifically, multiple statistical learning studies have provided evidence that a sensitivity to co-occurrences between inputs in many domains, including speech sounds, acoustic non-speech sounds, and visual objects (e.g., Bulf et al., 2011; Samuelson & Smith, 1999), emerges in infancy and persists into adulthood. Moreover, beyond being consistent with this prior evidence, the present findings build upon it by suggesting that sensitivity to co-occurrence regularities also contributes to the domain of semantic organization.

Second, the present findings corroborate evidence from numerous studies with children (e.g., Blaye et al., 2006; Lucariello et al., 1992; Walsh et al., 1993) and a handful of studies with adults (Lin & Murphy, 2001; Murphy, 2001) for the presence of links in semantic organization that may be derived from co-occurrence, such as schematic and thematic relatedness. Moreover, in contrast with schematic and thematic relatedness, which are constructs subjectively defined by researchers, the present findings highlight co-occurrence regularities as a measurable source of input in the environment that may shape semantic organization.

Contribution of Taxonomic Relations.—Taken together, the results of the three present experiments suggested that an influence of taxonomic relatedness came to supplement co-occurrence with development. Specifically, Experiments 1 and 2 did not detect significant effects of taxonomic relations at the group level in young children, and instead only detected weak and uncommon effects within individual children. Experiment 3 did detect an influence of taxonomic relations within young children as a group due to its use of a more sensitive, graded measure, but as in Experiments 1 and 2, this influence was weaker than the influence of co-occurrence. Across experiments, similar effects of taxonomic relations and co-occurrence were only observed in adults. Here, we evaluate this developmental trajectory in the context of prior research.

The degree to which taxonomic relations contribute to semantic organization at various points in development has been the subject of extensive prior research that has yielded conflicting findings. Numerous studies using a variety of behavioral paradigms have provided evidence that taxonomic relations only begin to contribute at the group level relatively late in development (e.g., Blaye et al., 2006; Lucariello et al., 1992; Tversky, 1985; Walsh et al., 1993), and a similarly large body of studies have provided evidence for

early taxonomic organization (e.g., Bauer & Mandler, 1989; Deák & Bauer, 1996; Gelman & Markman, 1986; Waxman & Namy, 1997). In spite of the contradiction between these bodies of research, we propose here that our present findings can be reconciled with both.

Evidence for Late Taxonomic Onset.: Results from several prior studies using a variety of behavioral paradigms that have observed an influence of taxonomic relatedness only in older children (e.g., age six and above), often following the earlier emergence of influences of relations that may be derived from co-occurrence or perceptual similarity. For example, many prior studies have investigated children's semantic organization using match-to-sample paradigms, in which participants are presented with a sample item (e.g., *dog*), and two choice items that are each related to the target in a different way (e.g., *elephant* and *bone*), and must select one choice item to match to the sample. Some studies that have used this approach have observed that, although older children may reliably choose taxonomic matches, young children do not (Lucariello et al., 1992; Tversky, 1985; Walsh et al., 1993). A similar pattern in which a robust influence of taxonomic relatedness is observed only in older children has emerged from studies that have inferred knowledge of semantic relations from sorting (Blaye et al., 2006), list recall (Bjorklund & Jacobs, 1985; Monnier & Bonthoux, 2011), and word association (Nelson, 1977)[5] paradigms. The present findings provide nuance to this apparent trajectory by suggesting that an influence of taxonomic relations is not entirely absent in young children, but is instead comparatively weak and uncommon, such that it is more readily detected when using a sensitive, graded measure such as the dynamics of looking behavior measured in Experiment 3.

Evidence for Early Taxonomic Onset.: Oher prior studies have yielded results that appear to demonstrate taxonomic knowledge that is detectable at the group level early in development. Specifically, in studies using variants of the match-to-sample paradigm conducted by Bauer and Mandler (1989); Deák and Bauer (1996); Gelman and Markman (1986); and Waxman and Namy (1997), young children consistently chose taxonomic matches, either throughout the study or under specific conditions. The evidence from our present experiments also suggests contributions of taxonomic relations in young children, and is only inconsistent with these prior findings in the strength and prevalence of these contributions.

One potential explanation for this difference in strength and prevalence of taxonomic knowledge is that additional information that could support taxonomic choices was available in prior studies showing strong, prevalent taxonomic influences. For example, in some of these studies, many target items are likely to have been visually similar to (e.g., *car* and *jeep*, *pot* and *skillet*) and/or co-occurring with (e.g., *chair* and *table*) their taxonomic matches. Moreover, in some of these studies, targets and taxonomic matches were given either identical labels, which may act as perceptual features that contribute to similarity in young children (Sloutsky & Fisher, 2004), or co-occurring labels (e.g., *puppy* and *dog*), such that

---

[5]One exception to this pattern in word association patterns is the tendency for even young children to produce taxonomic (or "paradigmatic") responses to number words, such as responding "two" when prompted with the word "one". However, our analyses of co-occurrence in child-directed speech measured from CHILDES corpora suggest that nouns for numbers one through ten frequently co-occur, rendering it unclear whether these responses are driven by co-occurrence, or an understanding that number words belong to the same category.

taxonomic choices could be based on co-occurrence (Fisher, 2010; Fisher, Matlen, & Godwin, 2011). Similarly, the availability of co-occurrence and/or perceptual similarity in addition to taxonomic relatedness also characterizes stimuli used in many studies of semantic knowledge in infants (Bergelson & Aslin, 2017; Styles & Plunkett, 2009; Willits, Wojcik, Seidenberg, & Saffran, 2013).

To the authors' knowledge, the only group-level evidence of an influence of taxonomic relatedness in young children in the absence of additional supportive information comes from one of two cued-recall paradigm experiments conducted by Blewitt and Toppino (1991). Specifically, Blewitt and Toppino found that recall accuracy in preschool-age children given pairs of unrelated words was exceeded by the accuracy of children given pairs of words that another sample of children had judged as co-occurring in both experiments, but was only also exceeded by the accuracy of children given taxonomically related words (referred to as "coordinate" pairs) in Experiment 2. Although the authors identified the lack of a taxonomic influence on accuracy in Experiment 1 as "spurious" (p. 311, Blewitt & Toppino, 1991), this inconsistency at least suggests that the taxonomic influence was less robust and evident only in some children, just as in the results of our experiments.

Finally, we note that the present findings rule out an alternative explanation that the apparent weakness of taxonomic relations in children was simply due to the possibility that the paradigms used in the present experiments were more sensitive to co-occurrence than taxonomic relations. Specifically, such a bias in the paradigms would have also led to the appearance of stronger co-occurrence versus taxonomic effects in adults. In contrast, we observed similar co-occurrence and taxonomic effects in adults.

Taken together, although the evidence available from prior research is sufficiently equivocal to fuel further debate, the evidence supporting the possibility that development typically involves an early-emerging role for co-occurrence that is increasingly supplemented by taxonomic relatedness is also sufficiently strong to highlight the importance of developing accounts that can explain this trajectory. This topic is discussed further in the following section.

Future Directions

The results of the present experiments were most consistent with the predictions of the mechanistic account proposed by Sloutsky et al. (2017). Specifically, in this account, sensitivity to co-occurrence regularities fosters the formation of semantic relations between both concepts whose referents or labels directly co-occur with each other (e.g., *fork* and *spaghetti*) and concepts whose referents or labels share patterns of co-occurrence (e.g., *spaghetti and pie*, which both co-occur with *fork*), which are in turn often taxonomically related (Asr et al., 2016; Cree & Armstrong, 2012; Huebner & Willits, 2018; Jones et al., 2015; Landauer & Dumais, 1997). According to this perspective, the earlier influence on semantic knowledge of co-occurrence versus taxonomic relatedness occurs because the latter is *derived from* the former. This prediction was corroborated by the developmental trajectory observed in the present experiments.

However, the core mechanisms proposed in Sloutsky et al.'s (2017) account, in which semantic relations between words are formed purely based on regularities with which they either directly co-occur or share each other's patterns of co-occurrence, remain largely unexplored in human learners. Specifically, prior research investigating this possibility is limited to only a handful of recent studies suggesting that toddlers and children form relations between words that directly co-occur (Matlen, Fisher, & Godwin, 2015; Wojcik & Saffran, 2015). Therefore, one key future direction highlighted by the present experiments is to test whether exposure to empirically manipulated linguistic input in which some pairs of words directly co-occur, and others share each other's patterns of co-occurrence, does indeed drive the formation of corresponding semantic relations in children and adults.

Conclusions

The present experiments provided evidence that word-word co-occurrence regularities captures relations between concepts in the semantic organization of both young children and adults. With development, co-occurrence was *supplemented* rather than *replaced* by taxonomic relatedness. These findings highlight importance of developing theoretical accounts of semantic development that incorporate a key role for co-occurrence regularities from early childhood onward.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Appendix A

Table A1

T-scores and Resnik similarities for pairs in the Co-Occur, Taxonomic, and Unrelated conditions in Experiments 1 & 2.

	Item 1	Item 2	t-score	Resnik
Co-Occur	bottle	baby	6.13	1.37
	foot	shoe	5.30	0.61
	brush	hair	15.75	1.37
	cup	juice	7.68	0.61
	cheese	mouse	2.97	0.61
	car	street	4.40	2.49
	soup	spoon	3.09	0.61
	milk	cow	4.53	0.61
	paper	pencil	6.61	0.61
Taxonomic	ball	puzzle	−3.65	7.98
	pig	bear	−1.91	5.61

	Item 1	Item 2	t-score	Resnik
	horse	bunny	−0.54	5.61
	carrot	banana	−0.19	6.35
	fork	bowl	0.21	8.10
	popcorn	fries	0.00	6.97
	airplane	boat	0.83	6.47
	sock	pajamas	0.00	5.87
	chicken	owl	−1.38	6.94
Unrelated	crayon	frog	0.00	1.37
	towel	bread	0.12	0.61
	blocks	cereal	0.00	0.61
	balloon	tree	0.00	1.37
	sheep	pancake	0.00	0.61
	pizza	lion	0.96	0.61
	fish	bed	−8.73	1.37
	duck	swing	0.00	1.37
	ice cream	bicycle	0.00	0.61

Note. T-scores for word pairs that never co-occurred within the 7-word window are undefined. Values for these pairs have therefore been entered as 0.00.

Appendix B

Table B1

T-scores and Resnik similarities for pairs in the Co-Occur, Taxonomic, and Unrelated conditions in Experiment 3.

Target	Co-Occur	t-score	Resnik	Taxonomic	t-score	Resnik	Unrelated	t-score	Resnik
Car	Street	4.40	2.49	Bicycle	0.71	6.31	Stick	−2.68	1.37
Nose	Tissue	26.33	0.61	Tongue	0.34	5.21		0.18	0.61
Airplane	Sky	3.94	0.61	Boat	0.83	6.47	Bathtub	0.00	3.45
Cheese	Mouse	2.97	0.61	Ice Cream	0.00	5.46		0.00	0.61
Horse	Cowboy	2.09	1.82	Frog	−0.11	5.20	Cloud	0.00	0.61
Pizza	Oven	6.38	0.61	Chocolate	0.00	5.46		0.00	0.61
Bread	Knife	3.89	0.61	Muffin	0.48	8.51	Button	0.04	0.61
Foot	Shoe	5.30	0.61	Head	−0.20	4.65		0.00	0.61
Fish	Water	5.65	0.61	Bird	1.17	5.20	Train	−1.29	1.37
Bed	Pillow	6.19	2.49	Table	−0.43	6.19		−1.62	3.45
Cow	Milk	4.53	0.61	Tiger	0.44	5.61	Book	−1.58	1.37
Leg	Pants	2.06	0.61	Finger	0.28	6.06		−2.94	0.61
Cereal	Breakfast	5.17	3.45	Pancake	0.00	6.58	Shovel	0.00	0.61
Monkey	Zoo	3.67	1.37	Squirrel	0.00	5.61		0.00	1.37
Soup	Spoon	3.09	0.61	Juice	0.01	5.46	Telephone	0.00	0.61
Coat	Zipper	2.75	2.49	Sweater	1.45	6.78		−0.70	2.49
Bottle	Baby	6.13	1.37	Bowl	0.00	6.95	Door	0.00	2.49

Target	Co-Occur	t-score	Resnik	Taxonomic	t-score	Resnik	Unrelated	t-score	Resnik
Apple	Tree	3.16	1.37	Grapes	0.00	8.00		−2.21	1.37
Corn	Chicken	3.22	0.61	Pumpkin	0.34	5.91	Star	0.00	0.61
Sock	Foot	4.93	0.61	Hat	0.54	5.87		0.00	1.37
Owl	Moon	5.25	1.37	Turtle	0.00	5.20	Nail	0.00	1.37
Bucket	Hose	5.99	2.49	Jar	0.45	5.32		0.00	3.45

Note. T-scores for word pairs that never co-occurred within the 7-word window are undefined. Values for these pairs have therefore been entered as 0.00

References

Princeton University. (2010). About WordNet. http://wordnet.princeton.edu

Arias-Trejo N, & Plunkett K (2009). Lexical–semantic priming effects during infancy. Philosophical Transactions of the Royal Society B: Biological Sciences, 364, 3633–3647.

Asr FT, Willits JA, & Jones MN (2016). Comparing predictive and co-occurrence based models of lexical semantics trained on child-directed speech Proceedings of the 38th Annual Meeting of the Cognitive Science Society Philadelphia, PA.

Baayen RH, Davidson DJ, & Bates DM (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language, 59, 390–412.

Bates D, Maechler M, Bolker B, & Walker S (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67, 1–48.

Bauer PJ, & Mandler JM (1989). Taxonomies and triads: Conceptual organization in one-to two-year-olds. Cognitive Psychology, 21, 156–184.

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

Bjorklund DF, & Jacobs JW (1985). Associative and categorical processes in children's memory: The role of automaticity in the development of organization in free recall. Journal of Experimental Child Psychology, 39, 599–617.

Blaye A, Bernard-Peyron V, Paour J-L, & Bonthoux F (2006). Categorical flexibility in children: Distinguishing response flexibility from conceptual flexibility. European Journal of Developmental Psychology, 3, 163–188.

Blewitt P, & Toppino TC (1991). The development of taxonomic structure in lexical memory. Journal of Experimental Child Psychology, 51, 296–319.

Bower GH, Clark MC, Lesgold AM, & Winzenz D (1969). Hierarchical retrieval schemes in recall of categorized word lists. Journal of Verbal Learning and Verbal Behavior, 8, 323–343.

Bulf H, Johnson SP, & Valenza E (2011). Visual statistical learning in the newborn infant. Cognition, 121, 127–132. [PubMed: 21745660]

Carey S (1985). Conceptual change in childhood. Cambridge, Massachusetts: MIT Press.

Cree GS, & Armstrong BC (2012). Computational models of semantic memory The Cambridge Handbook of Psycholinguistics (pp. 259–282). Cambridge: Cambridge University Press.

Deák G, & Bauer P (1996). The dynamics of preschoolers' categorization choices. Child Development, 67, 740–767. [PubMed: 8706524]

Fenson L, Vella D, & Kennedy M (1989). Children's knowledge of thematic and taxonomic relations at two years of age. Child Development, 60, 911–919. [PubMed: 2758885]

Fisher AV (2010). What's in the name? Or how rocks and stones are different from bunnies and rabbits. Journal of Experimental Child Psychology, 105, 198–212. [PubMed: 20005527]

Fisher AV, Matlen BJ, & Godwin KE (2011). Semantic similarity of labels and inductive generalization: Taking a second look. Cognition, 118, 432–438. [PubMed: 21227407]

Fox J, & Weisberg S (2011). An R Companion to Applied Regression (Second Edition ed.). Thousand Oaks, CA: Sage.

Frank MC, Braginsky M, Yurovsky D, & Marchman VA (2016). Wordbank: An open repository for developmental vocabulary data. Journal of Child Language.

Frermann L, & Lapata M (2015). Incremental Bayesian Category Learning From Natural Language. Cognitive Science, 40, 1333–1381. [PubMed: 26534863]

Fulkerson AL, & Waxman SR (2007). Words (but not tones) facilitate object categorization: Evidence from 6-and 12-month-olds. Cognition, 105, 218–228. [PubMed: 17064677]

Gellatly AR, & Gregg VH (1975). The effects of negative relatedness upon word-picture and word-word comparisons and subsequent recall. British Journal of Psychology, 66, 311–323. [PubMed: 1182400]

Gelman SA, & Coley JD (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. Developmental Psychology, 26, 796.

Gelman SA, & Markman EM (1986). Categories and induction in young children. Cognition, 23, 183–209. [PubMed: 3791915]

Goldberg RF, & Thompson-Schill SL (2009). Developmental "roots" in mature biological knowledge. Psychological Science, 20, 480–487. [PubMed: 19399979]

Heit E (2000). Properties of inductive reasoning. Psychonomic Bulletin & Review, 7, 569–592. [PubMed: 11206199]

Hofmann MJ, Biemann C, Westbury C, Murusidze M, Conrad M, & Jacobs AM (2018). Simple Co-Occurrence Statistics Reproducibly Predict Association Ratings. Cognitive Science, 42, 2287–2312. [PubMed: 30098213]

Howard DV, & Howard JH (1977). A multidimensional scaling analysis of the development of animal names. Developmental Psychology, 13, 108.

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

Huettig F, & Altmann GT (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. Cognition, 96, B23–B32. [PubMed: 15833303]

Huettig F, Quinlan PT, McDonald SA, & Altmann GT (2006). Models of high-dimensional semantic space predict language-mediated eye movements in the visual world. Acta Psychologica, 121, 65–80. [PubMed: 16098943]

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

Jimura K, Hirose S, Wada H, Yoshizawa Y, Imai Y, Akahane M, … Konishi S (2016). Relatedness-dependent rapid development of brain activity in anterior temporal cortex during pair-association retrieval. Neuroscience Letters, 627, 24–29. [PubMed: 27233220]

Jones MN, Willits J, & Dennis S (2015). Models of semantic memory In Busemeyer J & Townsend J (Eds.), Oxford Handbook of Mathematical and Computational Psychology (pp. 232–254). New York, NY: Oxford University Press.

Kemp C, & Tenenbaum JB (2008). The discovery of structural form. Proceedings of the National Academy of Sciences, 105, 10687–10692.

Klem M, Melby-Lervåg M, Hagtvet B, Lyster SAH, Gustafsson JE, & Hulme C (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. Developmental Science, 18, 146–154. [PubMed: 24986395]

Landauer TK, & Dumais ST (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychological Review, 104, 211.

Lin EL, & Murphy GL (2001). Thematic relations in adults' concepts. Journal of Experimental Psychology: General, 130, 3–28. [PubMed: 11293459]

Lucariello J, Kyratzis A, & Nelson K (1992). Taxonomic knowledge: What kind and when? Child Development, 63, 978–998.

MacWhinney B (2000). The CHILDES project: The database (Vol. 2): Psychology Press.

Matlen BJ, Fisher AV, & Godwin KE (2015). The influence of label co-occurrence and semantic similarity on children's inductive generalization. Frontiers in Psychology, 6, 1146. [PubMed: 26347672]

McClelland JL, & Rogers TT (2003). The parallel distributed processing approach to semantic cognition. Nature Reviews Neuroscience, 4, 310–322. [PubMed: 12671647]

Mirman D, Dixon JA, & Magnuson JS (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. Journal of Memory and Language, 59, 475–494. [PubMed: 19060958]

Mirman D, & Graziano KM (2012). Individual differences in the strength of taxonomic versus thematic relations. Journal of Experimental Psychology: General, 141, 601. [PubMed: 22201413]

Mirman D, & Magnuson JS (2009). Dynamics of activation of semantically similar concepts during spoken word recognition. Memory & Cognition, 37, 1026–1039. [PubMed: 19744941]

Monnier C, & Bonthoux F (2011). The semantic-similarity effect in children: Influence of long-term knowledge on verbal short-term memory. British Journal of Developmental Psychology, 29, 929–941. [PubMed: 21995745]

Murphy GL (2001). Causes of taxonomic sorting by adults: A test of the thematic-to-taxonomic shift. Psychonomic Bulletin & Review, 8, 834–839. [PubMed: 11848607]

Nelson K (1977). The syntagmatic-paradigmatic shift revisited: a review of research and theory. Psychological Bulletin, 84, 93. [PubMed: 322183]

Rosch E (1975). Basic objects in natural categories: Language Behavior Research Laboratory, University of California.

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

Ross BH, & Murphy GL (1999). Food for thought: Cross-classification and category organization in a complex real-world domain. Cognitive Psychology, 38, 495–553. [PubMed: 10334879]

Sadeghi Z, McClelland JL, & Hoffman P (2015). You shall know an object by the company it keeps: An investigation of semantic representations derived from object co-occurrence in visual scenes. Neuropsychologia, 76, 52–61. [PubMed: 25196838]

Samuelson LK, & Smith LB (1999). Early noun vocabularies: do ontology, category structure and syntax correspond? Cognition, 73, 1–33. [PubMed: 10536222]

Sloutsky VM (2010). From perceptual categories to concepts: What develops? Cognitive Science, 34, 1244–1286. [PubMed: 21116483]

Sloutsky VM, & Fisher AV (2004). Induction and categorization in young children: a similarity-based model. Journal of Experimental Psychology: General, 133, 166–187. [PubMed: 15149249]

Sloutsky VM, Yim H, Yao X, & Dennis S (2017). An associative account of the development of word learning. Cognitive Psychology, 97, 1–30. [PubMed: 28641208]

Smiley SS, & Brown AL (1979). Conceptual preference for thematic or taxonomic relations: A nonmonotonic age trend from preschool to old age. Journal of Experimental Child Psychology, 28, 249–257.

Smith LB, & Heise D (1992). Perceptual similarity and conceptual structure In Burns B (Ed.), Advances in Psychology: Percepts, Concepts, and Categories (Vol. 93, pp. 233–272). North-Holland: Elsevier.

Spence DP, & Owens KC (1990). Lexical co-occurrence and association strength. Journal of Psycholinguistic Research, 19, 317–330.

Storm C (1980). The semantic structure of animal terms: A developmental study. International Journal of Behavioral Development, 3, 381–407.

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

Tversky B (1985). Development of taxonomic organization of named and pictured categories. Developmental Psychology, 21, 1111–1119.

Walsh M, Richardson K, & Faulkner D (1993). Perceptual, thematic and taxonomic relations in children's mental representations: Responses to triads. European Journal of Psychology of Education, 8, 85–102.

Waxman SR, & Namy LL (1997). Challenging the notion of a thematic preference in young children. Developmental Psychology, 33, 555–567. [PubMed: 9149935]

Willits JA, Wojcik EH, Seidenberg MS, & Saffran JR (2013). Toddlers activate lexical semantic knowledge in the absence of visual referents: Evidence from auditory priming. Infancy, 18, 1053–1075.

Wisniewski EJ, & Bassok M (1999). What makes a man similar to a tie? Stimulus compatibility with comparison and integration. Cognitive Psychology, 39, 208–238. [PubMed: 10631012]

Wojcik EH, & Saffran JR (2015). Toddlers encode similarities among novel words from meaningful sentences. Cognition, 138, 10–20. [PubMed: 25704579]
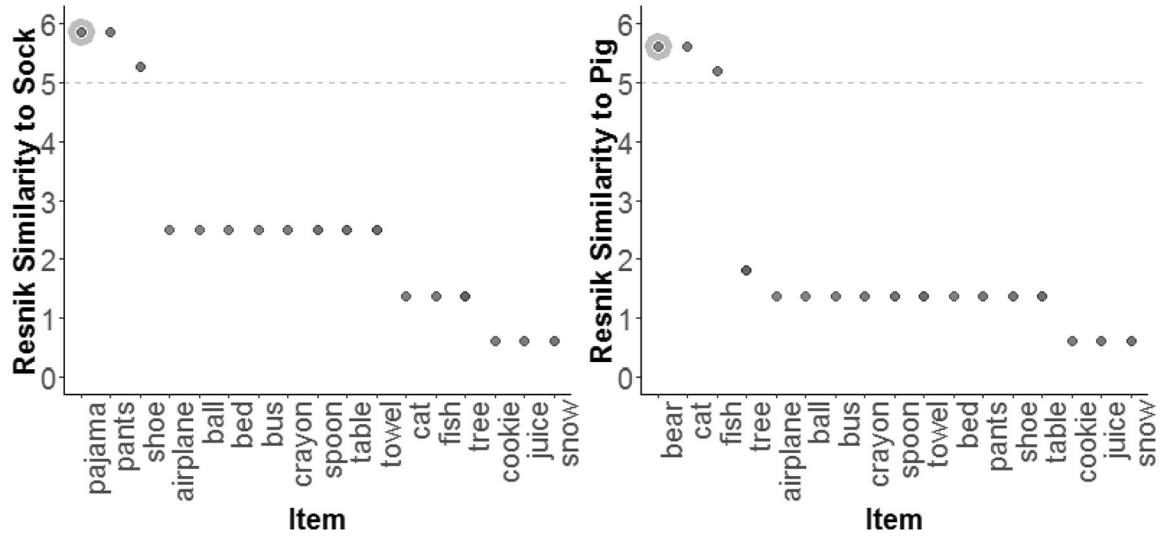
Figure 1.
Graphs depicting Resnik similarity between Taxonomic pairs versus other items. Each graph depicts the Resnik similarity between one item from a Taxonomic pair and: (1) The other item from the pair (highlighted), (2) Other items from the same taxonomic category, and (3) Items from other categories. These graphs depict that members of the same taxonomic category had Resnik similarities greater than five, whereas members of other categories had similarities substantially lower than five.
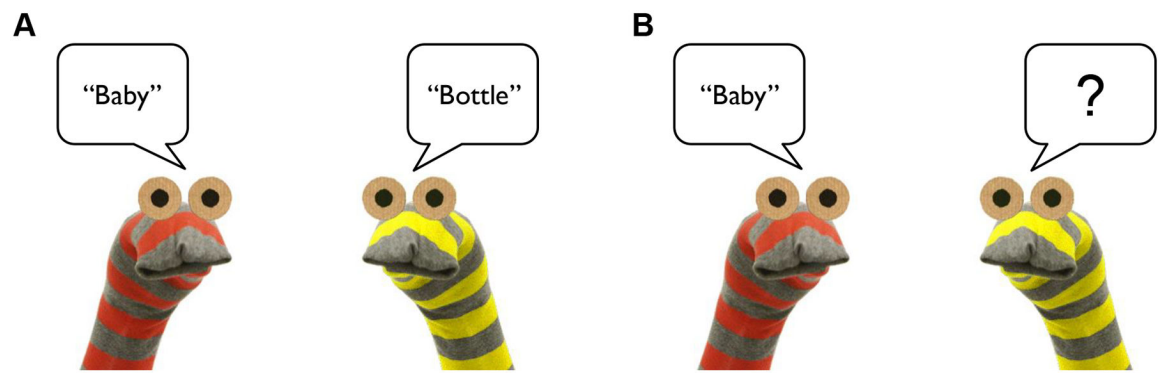
A   **B**

Figure 2.
Schematic of a trial in the Study Phase (A), in which one puppet "says" a Cue and the other a Target word, and a trial in the Test Phase (B), in which one puppet says the Cue, and the participant attempts to recall the Target.
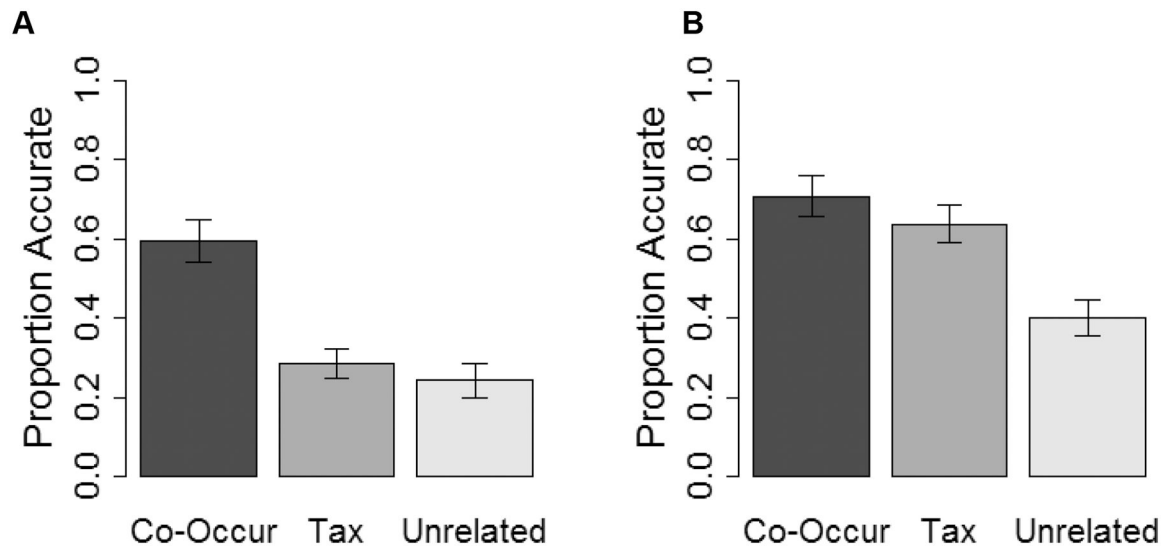
Figure 3.
Proportion accurate in children (panel A) and adults (panel B) in the three Relatedness conditions. Error bars represent standard errors of the means.
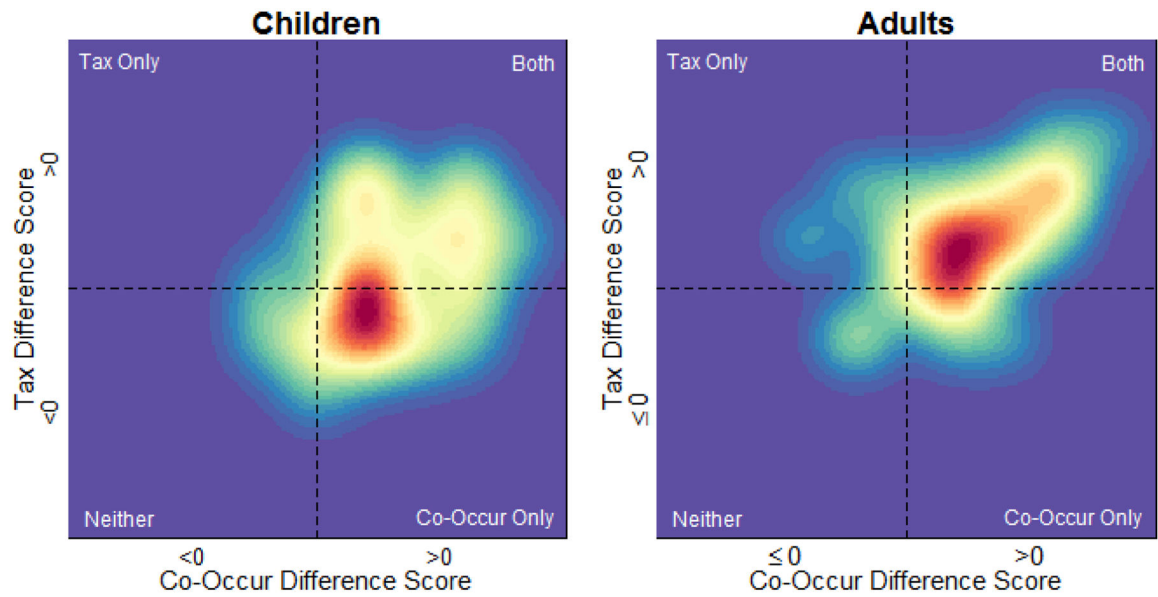
Figure 4.
Kernel densities for Co-Occur and Taxonomic Difference Scores in Experiment 1. Difference Scores were calculated by comparing the Co-Occur and Taxonomic conditions to the Unrelated condition, such that larger values correspond to larger influences of a given condition (i.e., greater improvement in accuracy).
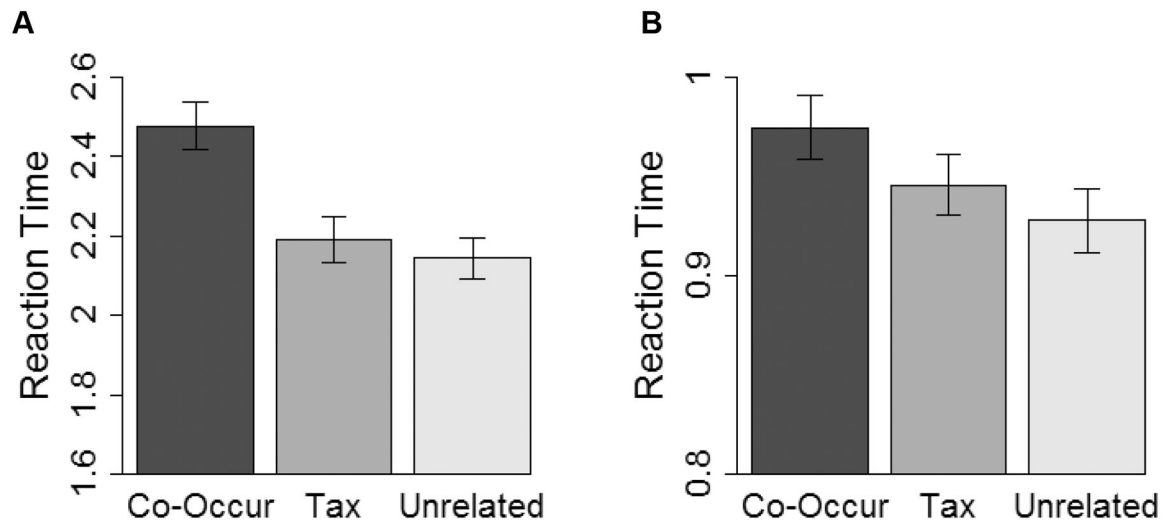
Figure 5.
Reaction Times in children (panel A) and adults (panel B) in the three Relatedness conditions. Error bars represent standard errors of the means. The y-axes for the two age groups are different because children's reaction times were substantially longer than those of adults.
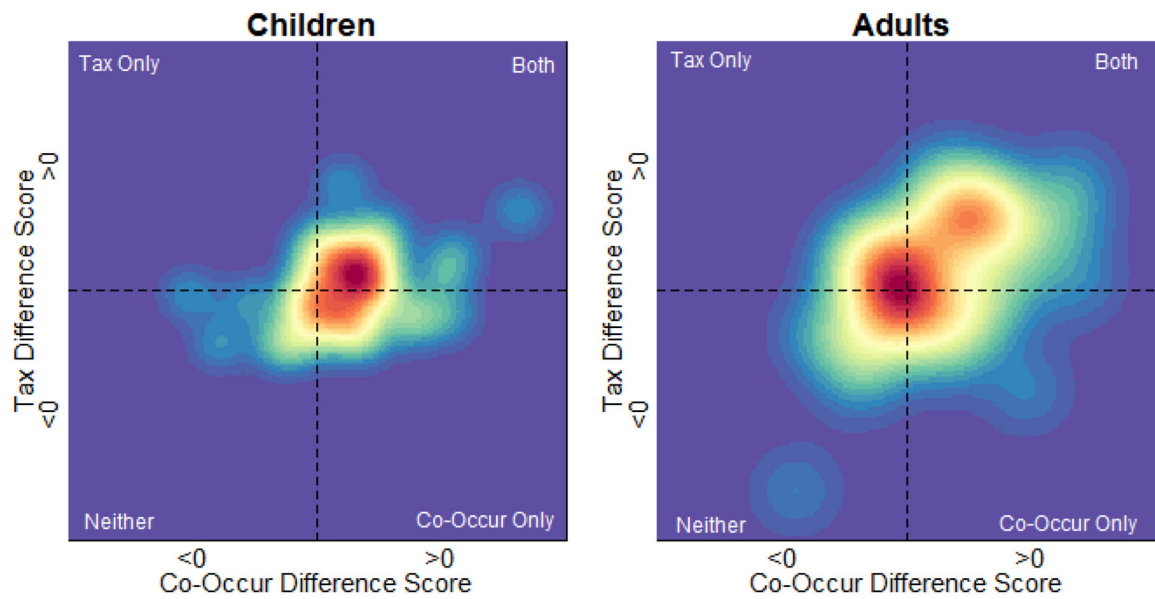
Figure 6.
Kernel densities for Co-Occur and Taxonomic Difference Scores in Experiment 2. Difference Scores were calculated by comparing the Co-Occur and Taxonomic conditions to the Unrelated condition, such that larger values correspond to larger influences of a given condition (i.e., greater slowing of reaction time).
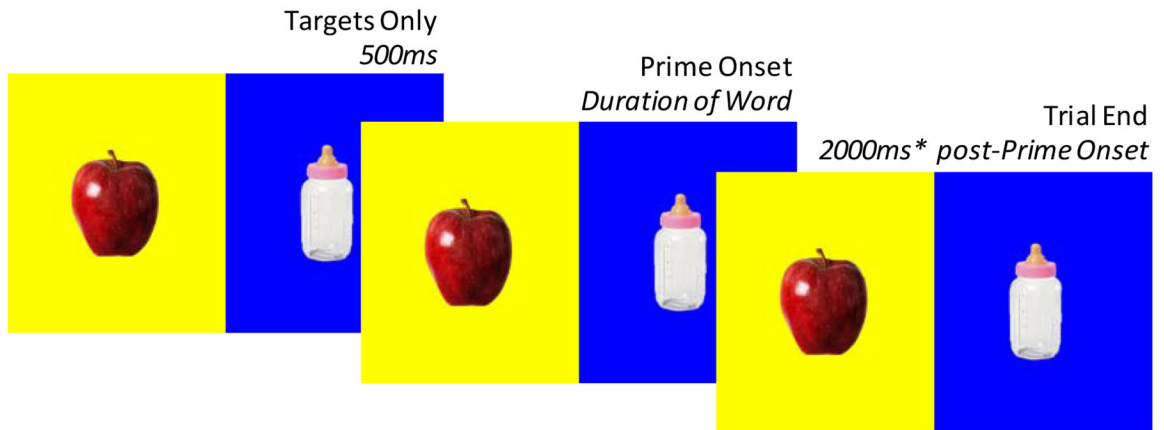
Figure 7.
Timing of events in Experiment 3 trials. Note: The trial ended 2000ms post-Prime Onset in Experimental trials only. In cover task filler trials, it ended either: When terminated by the experimenter upon observing the participant clicking the yellow or blue button, or after 5000ms.
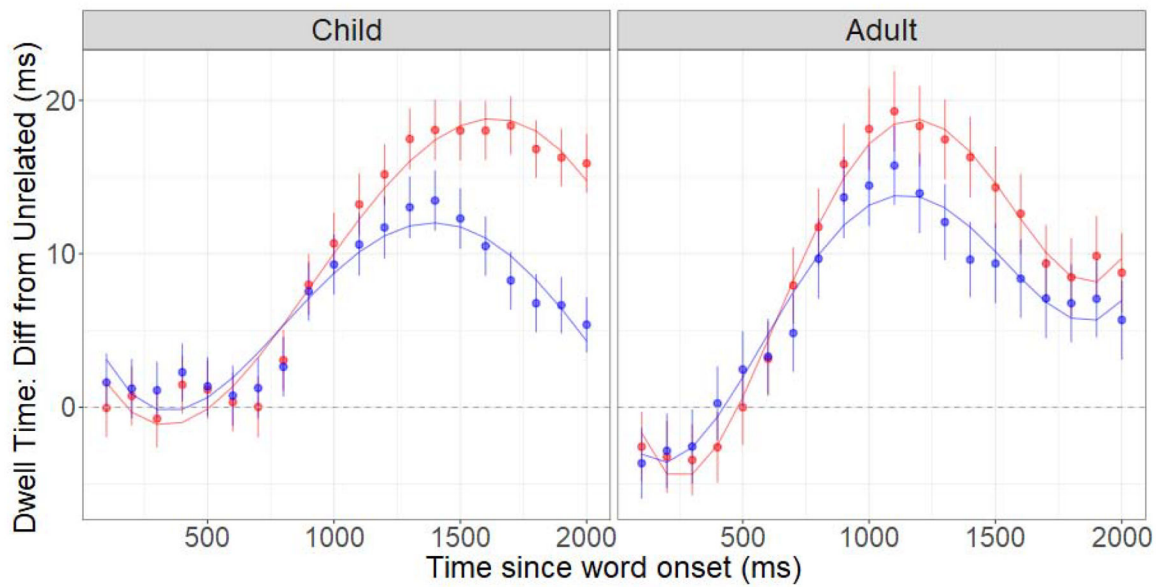
Figure 8.
Difference from Unrelated values in the Co-Occur (red) and Taxonomic (blue) conditions in Children and Adults, plotted with lines depicting the fitted values from the models. Error bars depict standard errors of the mean.

Table 1

Pairs of words used in the Co-Occur, Taxonomic, and Unrelated conditions in Experiments 1 & 2.

Co-Occur		Taxonomic		Unrelated	
Bottle	Baby	Ball	Puzzle	Crayon	Frog
Foot	Shoe	Pig	Bear	Towel	Bread
Brush	Hair	Horse	Bunny	Blocks	Cereal
Cup	Juice	Carrot	Banana	Balloon	Tree
Cheese	Mouse	Fork	Bowl	Sheep	Pancake
Car	Street	Popcorn	Fries	Pizza	Lion
Soup	Spoon	Airplane	Boat	Fish	Bed
Milk	Cow	Sock	Pajamas	Duck	Swing
Paper	Pencil	Chicken	Owl	Ice Cream	Bicycle

Note. Only 8 pairs from each condition were used in Experiment 1. A 9th pair was added to each condition (bottom row) in Experiment 2.

Table 2

Pair Sets in Experiment 3. Targets were presented as pictures, and Primes as words.

Pair Set	Target	Co-Occur Prime	Tax Prime	Unrelated Prime
1	Car	Street	Bicycle	Stick
	Nose	Tissue	Tongue	
2	Airplane	Sky	Boat	Bathtub
	Cheese	Mouse	Ice Cream	
3	Horse	Cowboy	Frog	Cloud
	Pizza	Oven	Chocolate	
4	Bread	Knife	Muffin	Button
	Foot	Shoe	Head	
5	Fish	Water	Bird	Train
	Bed	Pillow	Table	
6	Cow	Milk	Tiger	Book
	Leg	Pants	Finger	
7	Cereal	Breakfast	Pancake	Shovel
	Monkey	Zoo	Squirrel	
8	Soup	Spoon	Juice	Telephone
	Coat	Zipper	Sweater	
9	Bottle	Baby	Bowl	Door
	Apple	Tree	Grapes	
10	Corn	Chicken	Pumpkin	Star
	Sock	Foot	Hat	
11	Owl	Moon	Turtle	Nail
	Bucket	Hose	Jar	

Table 3

Results of growth curve analysis of Target Dwell Times. Parameter estimates are for the Co-Occur and Taxonomic conditions relative to the Unrelated condition. Non-significant parameter estimates are in italics.

Model Term	Age Group	Co-Occur		Taxonomic	
		Est. (SE)	*p*	Est. (SE)	*p*
Intercept	Child	9.599 (1.852)	<.001	6.389 (1.852)	<.001
Linear	Child	31.351 (5.603)	<.001	13.229 (5.603)	.020
Quadratic	Child	*−4.702 (5.057)*	*.355*	*−9.130 (5.057)*	*.075*
Intercept	Adult	8.986 (2.605)	<.001	6.768 (2.605)	.011
Linear	Adult	21.462 (7.149)	.003	14.384 (7.149)	.047
Quadratic	Adult	−23.151 (5.097)	<.001	−18.836 (5.097)	<.001

Table 4

Results of growth curve analysis of Difference from Unrelated. Parameter estimates are for the Co-Occur relative to the Taxonomic condition. Non-significant parameters are in italics.

Model Term	Age Group	Co-Occur versus Taxonomic Est. (SE)	p
Intercept	Child	*3.210 (1.970)*	*.113*
Linear	Child	18.122 (5.938)	.004
Quadratic	Child	*4.428 (5.094)*	*.390*
Intercept	Adult	*2.217 (2.493)*	*.378*
Linear	Adult	*7.078 (7.541)*	*.352*
Quadratic	Adult	*−4.315 (5.169)*	*.409*