



Published in final edited form as:

Psychol Sci. 2011 August ; 22(8): 1049–1057. doi:10.1177/0956797611413934.

Learning words in space and time: Probing the mechanisms behind the ‘suspicious coincidence’

John P. Spencer¹, Sammy Perone¹, Linda B. Smith², and Larissa K. Samuelson¹

¹Department of Psychology and Delta Center, University of Iowa

²Department of Psychological and Brain Sciences, Indiana University

Abstract

A major debate in the study of word learning centers on the extension of categories to new items. The rational approach assumes that learners make structured inferences about category membership, whereas the mechanistic approach emphasizes the attentional and memorial processes upon which generalization behaviors are based. Recent support for the rational view comes from a phenomenon called the “suspicious coincidence”: people generalize category membership narrowly when presented with three subordinate examples that share the same label and broadly when presented with one example. Across three experiments, we examine the mechanistic basis of the suspicious coincidence. Results show that the presentation of multiple subordinate examples only leads to narrow generalization when the exemplars are presented simultaneously, even when the number of examples is increased from three to six. These data demonstrate that the suspicious coincidence is firmly grounded in the general cognitive processes of attention, memory, and visual comparison.

Explanations of word learning can be broken into two broad classes—rational and mechanistic. In the rational approach, the problem of word learning is viewed as one of rational inference and described in terms of hypotheses, evidence, and structured inference (e.g., Chater & Manning, 2006; Heibeck & Markman, 1987; Xu & Tenenbaum, 2007a). In the mechanistic approach, word learning is viewed as the product of memorial and attentional processes and the dynamic representational states they engender (e.g., Gershkoff-Stowe & Hahn, 2007; Plunkett, 1997; Regier, 2003; Smith & Samuelson, 2006; Stokes & Klee, 2009). The relation between rational and mechanistic accounts is unresolved and controversial (McClelland et al., 2010; Griffiths et al., 2010) with some arguing that the two approaches are complementary rather than directly competitive (see Regier, 2003; Sakamoto, Jones & Love, 2008), and others claiming that these are explanations at different levels of analysis with different goals (Kemp & Tenenbaum, 2009). We will not resolve this controversy here, but add to the discussion by offering a mechanistic account of a new phenomenon called the *suspicious coincidence* that was predicted by Xu and Tenenbaum’s (2007a) rational account of structured probabilistic inference.

The phenomenon is this: A learner sees a big white dog with black spots and hears, “Look at the pretty *fep!*” Next, the learner sees another white dog with black spots and hears, “Look, another *fep!*” The learner then sees a person walking two white dogs with black spots and hears, “Two more *feps!*” The predicted inference is that the learner will infer that ‘fep’ refers to this specific type of white dog with black spots and will not infer that ‘fep’ refers to the class of dogs. To quote Xu and Tenenbaum (2007a), “Intuitively, this inference appears to be based on a *suspicious coincidence*: It would be quite surprising to observe only

Dalmatians called *feps* if in fact the word referred to all dogs and if the first four examples were a random sample of *feps* in the world” (p. 249).

Xu and Tenenbaum (2007a) reported that both children and adults make this inference when learning the referent of novel words. In particular, when shown three Dalmatians and given a novel name—*fep*—participants only generalized the name to other Dalmatians. By contrast, when participants were shown a single item—a single Dalmatian—and given a novel name—*fep*—participants generalized the name to all variety of dogs. This might seem a surprising result: Learners show broader generalization from just one instance (one Dalmatian) than from more instances (three nearly identical Dalmatians). Moreover, the result provides strong evidence for a learning system that represents probability distributions and makes structured inferences over those distributions. Xu and Tenenbaum (2007a) concluded that “an intuitive sensitivity to these sorts of suspicious coincidences is a core capacity enabling rapid word learning, and we argue it is best explained within a Bayesian framework” (p. 249).

The striking finding that led them to this conclusion—broader generalization from a single instance than from three (nearly identical) instances—is also consistent with mechanistic accounts couched in terms of memories and representations for learning events. This proposal takes as its starting point a hundred years of research in experimental psychology indicating that the breadth of generalization depends on the diversity of instances, the diversity of the contexts in which those instances are experienced, whether they are presented simultaneously versus sequentially (e.g., Gibson, 1969; Hahn, Bailey & Elvin, 2005; Honig & Day, 1962; James, 1890; MacIntosh, 1965; Rips & Collins, 1993), and the order with which instances are encountered (Garner, 1974; Medin & Bettger, 1994; Samuelson & Horst, 2007; Sandhofer & Doumas, 2008; Schyns, Goldstone & Thibaut, 1998). In the present case, two such task factors may be particularly critical: the fact that the exemplars are simultaneously visible in the task space and that they are nearly identical instances in close spatial proximity. Such simultaneous presentation in close proximity is expected to increase discrimination, fine-grained comparison, and memory for the specific shared features (Garner, 1974; Gentner & Namy, 2006; Gibson, 1969; Samuelson, Schutte, & Horst, 2009; Sandhofer & Smith, 2001). The presentation of a single instance by itself with no comparison may be expected to yield less focused representations and, thus, broader generalization (see, for example, Garner, 1974; Gibson, 1969; Honig & Day, 1962).

Accordingly, the present experiments used Xu and Tenenbaum’s task to examine the narrowness and breadth of generalization across manipulations of the spatial-temporal details of the experienced instances. We show that the suspicious coincidence depends critically on whether the instances are simultaneously visible in the task space, a result expected from mechanistic accounts given that the processes that perceive, encode, and remember events are spatially and temporally extended. This dependence on simultaneous presentation is evident across two experiments, even when the number of exemplars is increased from three to six. This provides a critical test because Bayesian accounts predict that category extensions should become narrower as the number of subordinate exemplars is increased.

General Method

Participants

There were 58 adult participants (19 in Experiments 1 and 3 and 20 in Experiment 2). Participants received credit in a college course or monetary compensation for their participation. All participants reported normal vision and gave informed consent.

Stimuli

The stimulus set matched that used by Xu and Tenenbaum (2007a). The stimuli consisted of 45 digitized color photographs of real objects distributed across the vegetable, vehicle, and animal domains. Seven objects were designated for labeling in each domain: there was one singleton example, three subordinate examples, three basic-level examples, and three superordinate examples (see Figure 1). As Figure 1 shows, the one singleton example was used as an example at all levels (see Xu & Tenenbaum, 2007a). Eight objects were designated for generalization in each domain: two subordinate, two basic-level, and four superordinate objects (Figure 1). Twelve nonsense words (e.g., foo) were paired with the examples across presentations. The mapping of words to examples was randomly assigned across participants.

Design and Procedure

Following Xu and Tenenbaum (2007a), participants were told that they were playing a game with Mr. Frog. Their task was to help Mr. Frog find the objects he wanted. At the beginning of each trial, participants were presented with a label for one example (e.g., Here is a fep) or several examples (e.g., Here are three feps) of a category, and asked to help Mr. Frog ‘find the other feps’ (Figure 2). The 24 generalization objects were randomly distributed across a 4 × 6 grid on the computer screen above the category examples. Participants selected their choice of objects that matched the example(s) by clicking on a picture using the mouse cursor. Participants clicked a “done” button when they had finished their selections.

For each domain, there was one trial for each of the four different example types (see Figure 1). These trials were blocked by example type across domains; thus, participants completed all trials of one type (e.g., the one example trial for animals, vegetables, and vehicles) before moving on to the next example type (e.g., the basic-level examples trial for animals, vegetables, and vehicles). As we discuss below, the first block of trials always involved either the one example trials or the three subordinate examples trials from each domain. The remaining blocks of trials were randomly ordered for each participant¹.

Method of Analysis

The suspicious coincidence is about how people generalize novel words in two critical conditions: when a single exemplar is presented versus when three subordinate examples are presented. This phenomenon is also primarily about differential generalization at one specific level—the basic level. In particular, when Xu and Tenenbaum (2007a) presented participants with a single exemplar (a Dalmatian), people generalized this item to roughly 57% of the basic-level matches (averaged across children and adults). By contrast, when presented with three subordinate examples (three Dalmatians), people generalized these items to roughly 7% of the basic-level matches. Responding at the other levels—subordinate and superordinate—were comparable across these two conditions (participants generalized to roughly 95% of the subordinate matches and roughly 7% of the superordinate matches).

Although we included all conditions to fully replicate Xu and Tenenbaum’s experiments, we focus our analyses on the changes in basic-level responding across experiments in two central conditions. Analyses of the remainder of the full data set (including the supplemental experiments), which is reported in Table 1, revealed no significant differences in responding

¹Xu and Tenenbaum (2007) did not block trials for the three exemplar trials; rather, they presented these trials in a pseudorandom order, counterbalanced across participants. These researchers also only used nine distinct words as labels for the twelve exemplar sets, using the same labels to refer to the one example and three subordinate example sets. Unfortunately, these details were not in the original report; they came to light in the review of this manuscript. We regret this inconsistency across studies; however, our three replications of the key findings (Experiment 1, Supplemental Experiments 1 and 2) demonstrate that these details are not critical to the suspicious coincidence.

across experiments. As in Xu and Tenenbaum (2007a), participants generalized to a high percentage of the subordinate matches across all exemplar conditions and experiments. An ANOVA comparing the percent of subordinate matches selected across example types (one example, three subordinates, three basic-levels, and three superordinates) and experiments (1-3, S1, S2) revealed no significant effects.

Table 1 also shows that, as in Xu and Tenenbaum (2007a), participants rarely selected superordinate matches in the one example and three subordinate examples conditions. The selection of superordinate matches increased when three basic-level examples were presented, and increased even further when three superordinate-level examples were presented. Critically, an ANOVA comparing superordinate responding across experiments showed no significant differences across experiments. Finally, we examined generalization at the basic-level when three basic-level and three superordinate-level examples were presented. When participants saw three basic-level examples, they generalized at the basic-level more than when they saw three superordinate examples. Once again, an ANOVA comparing performance across experiments revealed no significant differences in basic-level responding for these conditions across experiments.

Experiment 1

This experiment replicated Experiment 1 from Xu and Tenenbaum (2007a). As shown in Figure 3, and consistent with Xu and Tenenbaum's previous findings, participants selected a significantly higher percentage of basic-level matches when they saw one example than when they saw three subordinate examples, $t(18)=3.40$, $p < .01$, two-tailed. The magnitude of basic-level responding for the one example condition was comparable to the overall mean performance in Xu and Tenenbaum (2007a), but it was higher than children's basic-level performance and lower than adult's basic-level performance (see Figure 3). It is not clear why adults in our study were more conservative in their responding at the basic level relative to the adults in Xu and Tenenbaum. Nevertheless, our results robustly replicate the suspicious coincidence: participants showed significantly narrower generalization when three subordinate examples (e.g., three Dalmatians) were presented relative to when a single example (e.g., one Dalmatian) was presented.

Experiment 2

In Experiment 1, the exemplars were simultaneously in view and in close spatial proximity which is known to foster comparison, attention to shared features, and fine-grained discrimination (e.g., Gentner & Namy, 2006; Gibson, 1969; Samuelson et al., 2009). In Experiment 2, we manipulated this aspect of the task, presenting the three subordinate examples sequentially at three different locations. This sequential presentation is arguably more consistent with experienced instances in everyday word learning which are, as in our opening Dalmatian example, not usually experienced at the same time or place.

In this sequential version of the task, the first example was presented at the left location for 1 s and then removed, then the second example was presented at the middle location for 1 s and removed, and then the third example was presented at the right location for 1 s and removed. This sequence was repeated 2 times prior to the presentation of the generalization set for a total of 6 s of study time (see Figure 4). The exemplar display continued to loop through the three subordinate exemplars continuously while participants were making their selections until the participant clicked 'done'. Note that we presented the three subordinate examples in the first block of trials to ensure that performance on these critical trials was not influenced by generalizations on the other trials. Supplemental experiments 1 and 2 confirm that this manipulation in isolation does not modify the robustness of the suspicious coincidence, nor does the addition of an initial study phase.

Results and Discussion

As can be seen in Figure 3, manipulating the manner of presentation of the three subordinate exemplars had a dramatic impact on participants' performance—it reversed the *suspicious coincidence*. In particular, participants generalized more broadly at the basic level when they saw three subordinate examples in a sequence than when they saw one example in isolation, $t(19)=-3.01, p<.01$, two-tailed. The broader generalization in the three subordinate examples condition differed significantly from performance in the replication of Xu and Tenenbaum in Experiment 1, $t(37)=-4.28, p<.0001$, two-tailed. By contrast, the slightly narrower generalization in the one example condition did not differ from performance in Experiment 1, $t(37)=1.40, p = .17$, two-tailed. These results are consistent with a large literature on comparison processes (e.g., Gentner & Namy, 2006, see also, Vlach, Sandhofer & Kornell, 2008). They suggest that the phenomenon called the suspicious coincidence is critically linked to these processes and is sensitive to the space-time details of experience.

Experiment 3

The previous experiment showed that people generalize novel words differently when the exact same stimuli are presented sequentially at three locations versus simultaneously, a result consistent with a large literature on the effects of simultaneous and sequential presentation on perceptual learning (e.g., Gibson, 1969), attention (e.g., Chun & Nakayama, 2000), and relational and category reasoning (Gentner & Namy, 2006). One might argue, however, for an alternative interpretation that is perhaps more in line with statistical inference. At the limit of sequential presentation (e.g., with long temporal delays between presentations as one might experience in everyday life), the instances might not be remembered in detail or identified as multiple, different exemplars of the same subordinate group. This raises the question: did participants detect that multiple, different exemplars of the same type of thing were present? If not, they might have generalized broadly in Experiment 2 because they interpreted the display as containing only a single item (even though the display said 'here are three feps'). Informal discussions with participants after the testing session suggested that they knew, for instance, that the dogs were different examples of the same type of dog; nonetheless, we provide direct evidence on the issue in this experiment.

Another way to vary presentation to highlight commonalities and differences among similar items is to present the items sequentially, but all at the same location. When sequential items are presented in the same location with no delay in-between presentations, differences are easy to detect because visual attention does not have to shift to a new location and one can rely on lower-level sensory memory (see, e.g., Phillips, 1974). Thus, in this experiment, we showed exemplars sequentially at a single location to facilitate the task of identifying that multiple, different exemplars from the same subordinate group were being presented.

In addition to this manipulation, we made a second modification to directly pit the mechanistic and rational accounts. According to the 'size' principle in Xu and Tenenbaum's account (2007a), subordinate level hypotheses will be assigned greater probability than hypotheses at other hierarchical levels, and such differences are amplified *exponentially* as the number of consistent examples increases. For instance, with three examples as in the previous experiments, the likelihood ratio of subordinate and basic level hypotheses is inversely proportional to the ratio of their sizes, raised to the third power. Accordingly, we increased the number of examples in the present experiment to six. If Xu and Tenenbaum's analysis is right, the suspicious coincidence should be more dramatic and generalization should be very narrow. By contrast, if the suspicious coincidence depends principally on the type of detailed comparison afforded by simultaneous presentation, then, given the sequential presentation here, we should once again fail to see the suspicious coincidence.

To summarize, Experiment 3 was identical to the previous experiment except we presented six subordinate examples, these examples were presented for 1 s at a time at a single central location, and the text on the screen said that there were ‘six feps’ instead of ‘three feps’. As before, participants were given 6 s of study time before the generalization set appeared, and the sequence of exemplars looped continuously until the trial was done. The central question was whether participants would generalize narrowly based on the suspicious coincidence that six Dalmatians were all given the same novel label, or generalize broadly given the challenges of remembering and comparing multiple subordinate items presented sequentially.

Results and Discussion

As can be seen in Figure 3, the results replicated the key finding from Experiment 2—*sequential presentation leads to broad generalization at the basic level, even when six exemplars are presented*. Indeed, participants generalized broadly across *both* critical conditions with no significant difference between the one example and six subordinate examples conditions, $t(18)=-1.30, p = .21$. The broad generalization with six subordinate examples differed significantly from the three subordinate examples condition in Experiment 1, $t(36) = -3.64, p < .001$, two-tailed. The broad generalization with one example was comparable to performance in the same condition from Experiment 1, $t(37)=-.78, p = .44$, two-tailed. Thus, even when participants see six subordinate examples all given the same label, they still generalize broadly at the basic level if the examples are presented sequentially. Across experiments, then, *participants show a striking inability to detect the suspicious coincidence when exemplars are presented sequentially*. This shows that the spatial and temporal details of experience matter critically to the suspicious coincidence. Indeed, these details had a more profound influence on performance than the number of exemplars, suggesting that space and time might be more constraining of performance than the size principle proposed by Xu and Tenenbaum (2007a).

General Discussion

Outside of the laboratory, people (including young children) experience objects and hear labels distributed in space and time, across many different locations and many different times. Thus, understanding how the spatial and temporal dynamics of experienced instances influence word learning is important in its own right. The suspicious coincidence reported by Xu and Tenenbaum (2007a) is also important because it illustrates a core principle of statistical learning and thus provides support for a specific Bayesian model of probabilistic structured inference. As reported by Xu and Tenenbaum (2007a), the suspicious coincidence is not consistent with hypothesis elimination approaches (Berwick, 1986; Pinker, 1984; Siskind, 1996), nor is it consistent with connectionist / associative accounts (Colunga & Smith, 2005; Gasser & Smith, 1998; Regier, 1996; Roy & Pentland, 2004). Xu and Tenenbaum’s Bayesian model—a rational approach to word learning—is currently the only formal theory that has quantitatively captured this experimental result.

The experiments here considered the suspicious coincidence in the context of more basic perceptual and cognitive processes and prior experimental results on simultaneous and sequential presentations of instances in learning experiments. Results of Experiments 2 and 3 indicate that people do not generalize narrowly when multiple subordinate examples are presented sequentially, even when twice as many subordinate examples are presented. A critical question is why. We contend the answer lies in how participants compare items and extract similarities and differences with these different modes of presentation. Simultaneous presentation invites an emphasis on fine-grained similarities and differences. Participants might, for instance, notice that the green peppers from the exemplar set in Figure 1 all have a dark green region near the stem, leading them to select other green peppers with this same

dark green region in the generalization set (the subordinate matches), but reject peppers that do not share this feature (the basic-level matches). Sequential presentation, by contrast, affords a more global interpretation of similarity (see Samuelson, Schutte & Horst, 2009), leading participants to emphasize the overall light green shade of the peppers in the exemplar set. In this case, they would select the two subordinate matches and also the yellow-green pepper in the basic level set.

To our knowledge, there are five experiments showing the suspicious coincidence effect: Xu and Tenenbaum (2007a; Experiments 2 and 3); Xu and Tenenbaum (2007b; Experiment 1); and Gweon et al (2009; Experiments 1 and 2). In all of these cases, the exemplars were simultaneously visible in the task space in close spatial proximity. If the narrow generalization effect in these studies depends on the fine-grained comparisons afforded by simultaneous presentation, then the meaning of the suspicious coincidence itself may have to be rethought. At the very least, our data limit the applicability of the Bayesian model to situations in which instances are in close temporal and spatial proximity. This is a critical boundary condition if the goal is to explain real world word learning: the simultaneous experience of multiple instances of a subordinate category in close proximity is not the typical word-learning context.

There is no space or time in Xu and Tenenbaum's model. Thus, the perceptual, attentional and memory processes that led us to test the effects of sequential presentation are outside the purview of that account. Nevertheless, their model is designed to provide insight into word learning, and space and time matter to word learning. Experiment 3 did manipulate a factor that the Bayesian account claims is central to category extension—the number of exemplars. The findings, here, directly contradict the model's claim that the likelihood of generalizing at a particular level is scaled exponentially by the number of examples at that level. Sequential presentation effectively trumps the number of subordinate exemplars.

One question is whether the Bayesian model can be modified to capture our results. For instance, Xu and Tenenbaum (2007b) reported broader generalization at the basic level in a “learner-driven” condition where participants—rather than a knowledgeable teacher—selected a set of subordinate examples. This was predicted based on the assumption that learners treat items selected by a teacher as independent samples from the real extensions of novel words, while their own selections are uninformative about the extensions of novel words. Is it possible, then, that sequential presentation led participants in our study to treat the samples as biased or non-independent, yielding broader basic-level generalization? We know of no reason why participants might do this. Moreover, the Bayesian model predicts a drop in subordinate level responding when independence is not assumed (Xu & Tenenbaum, 2007b). This did not occur here (see Table 1). Thus, it is not clear how the Bayesian model might explain our results.

The perceptual and memory factors that motivated this study do not offer straightforward explanations of other findings that were directly predicted by the Bayesian account, for example, the role of the distribution of the generalization set (Gweon et al, 2009) and differential generalization in “teacher-driven” versus “learner-driven” conditions (Xu & Tenenbaum, 2007b). This brings us to the door step of the current controversy over rational and mechanistic accounts of word learning and categorization. Clearly, rational, structured inference captures something real about human cognition; just as clearly, human cognition is grounded in the real time dynamic processes of memory and attention at another level of description.

All of this underscores the utility of thinking about Marr's (1982) levels of analysis, not as separate levels, but as mutually-informative and reciprocal (see Sakamoto et al., 2008). In

the end, computational level accounts will be most impactful (and enduring as explanations) when they connect to other levels and the details of behavior in substantive ways. In this context, the present data present a challenge to one particular Bayesian account and show that it must move to greater specificity with respect to the role of space and time in learning from instances. There are a large number of phenomena in experimental psychology—such as the effects of simultaneous versus sequential comparison—that are both principle-like in their robustness across experimental tasks and principally about the dynamic properties of memory, attention, and generalization (Samuelson & Smith, 2000). Our results suggest that properties of statistical learning and outcomes such as the suspicious coincidence are critically connected to these more general processes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Eliot Hazeltine for help programming the basic task. We also thank Gavin Jenkins, Joshua Tenenbaum, and an anonymous reviewer for useful comments on earlier versions of this manuscript. This work supported by NIHMH R01 MH 62480 and NSF HSD 0527698 awarded to JPS, and NICHD R01 HD 045713 awarded to LKS.

References

- Berwick, RC. Learning from positive-only examples: The subset principle and three case studies. In: Carbonell, JG.; Michalski, RR.; Mitchell, TM., editors. *Machine Learning: An artificial intelligence approach*. Vol. 2. Los Altos, CA: Morgan Kaufman; 1986. p. 625-645.
- Chater N, Manning CD. Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences Special Issue: Probabilistic Models of Cognition*. 2006; 10(7):335–344.
- Chun MM, Nakayama K. On the functional role of implicit visual memory for the adaptive deployment of attention across scenes. *Visual Cognition. Special Issue: Change Blindness and Visual Memory*. 2000; 7(1-3):65–81.
- Colunga E, Smith LB. From the lexicon to expectations about kinds: A role for associative learning. *Psychological Review*. 2005; 112:347–382. [PubMed: 15783290]
- Garner, WR. *The processing of information and structure*. Potomac, MD: Erlbaum; 1974.
- Gasser M, Smith LB. Learning nouns and adjectives: A connectionist approach. *Language and Cognitive Processes*. 1998; 13:269–306.
- Gentner D, Namy LL. Analogical processes in language learning. *Current Directions in Psychological Science*. 2006; 15(6):297–301.
- Gershkoff-Stowe L, Hahn ER. Fast mapping skills in the developing lexicon. *Journal of Speech, Language, and Hearing Research*. 2007; 50(3):682–696.
- Gibson, EJ. *Principles of perceptual learning and development*. East Norwalk, CT, US: Appleton-Century-Crofts; 1969.
- Griffiths TL, Chater N, Kemp C, Perfors A, Tenenbaum JB. Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Science*. 2010; 14:357–364.
- Gweon H, Tenenbaum JB, Schulz L. What are you trying to tell me? A Bayesian model of how toddlers can simultaneously infer property extension and sampling processes. *Proceedings of the 31st Annual Conference of the Cognitive Science Society*. 2009:1282–1287.
- Hahn U, Bailey TM, Elvin LBC. Effects of category diversity on learning, memory, and generalization. *Memory & Cognition*. 2005; 33(2):289–302.
- Heibeck TH, Markman EM. Word learning in children: An examination of fast mapping. *Child Development*. 1987; 58(4):1021–1034. [PubMed: 3608655]
- Honig WK, Day RW. Discrimination and generalization on a dimension of stimulus difference. *Science*. 1962; 138:29–31. [PubMed: 13908709]
- James, W. *The principles of psychology*. Henry Holt & Co; New York: 1890.

- Kemp C, Tenenbaum JB. Structured statistical models of inductive reasoning. *Psychological Review*. 2009; 116(1):20–58. [PubMed: 19159147]
- MackIntosh NJ. The effect of attention on the slope of generalization gradients. *British Journal of Psychology*. 1965; 56(1):87–93. [PubMed: 14292058]
- Marr, D. *Vision A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: W.H. Freeman; 1982.
- McClelland JL, Botvinick MM, Noelle DC, Plaut DC, Rogers TT, Seidenberg MS, Smith LB. Letting structure emerge: Connectionist and dynamical systems approaches to understanding cognition. *Trends in Cognitive Science*. 2010; 14:348–356.
- Medin DL, Bettger JG. Presentation order and recognition of categorically related examples. *Psychonomic Bulletin & Review*. 1994; 1(2):250–254.
- Phillips WA. On the distinction between sensory storage and short-term visual memory. *Percept Psychophys*. 1974; 16:283–290.
- Pinker, S. *Language learnability and language development*. Cambridge, MA: Harvard University Press; 1984.
- Plunkett K. Theories of early language acquisition. *Trends in Cognitive Sciences*. 1997; 1(4):146–153. [PubMed: 21223888]
- Regier, T. *The human semantic potential: Spatial language and constrained connectionism*. Cambridge, MA: MIT Press; 1996.
- Regier T. Emergent constraints on word-learning: A computational perspective. *Trends in Cognitive Sciences*. 2003; 7(6):263–268. [PubMed: 12804693]
- Rips LJ, Collins A. Categories and resemblance. *Journal of Experimental Psychology: General*. 1993; 122(4):468–486. [PubMed: 8263465]
- Roy D, Pentland A. Learning words from sights and sounds: A computational model. *Cognitive Science*. 2004; 26:113–146.
- Sakamoto Y, Jones M, Love BC. Putting the psychology back into psychological models: Mechanistic versus rational approaches. *Memory & Cognition*. 2008; 36(6):1057–1065.
- Samuelson LK, Horst JS. Dynamic noun generalization: Moment-to-moment interactions shape children's naming biases. *Infancy*. 2007; 11:97–110.
- Samuelson LK, Schutte AR, Horst JS. The dynamic nature of knowledge: Insights from a dynamic field model of children's novel noun generalizations. *Cognition*. 2009; 110:322–345. [PubMed: 19131050]
- Samuelson LK, Smith LB. Grounding development in cognitive processes. *Child Development*. 2000; 71(1):98–106. [PubMed: 10836563]
- Sandhofer CM, Smith LB. Why children learn color and size words so differently: Evidence from adults' learning of artificial terms. *Journal of Experimental Psychology: General*. 2001; 130(4):600–617. [PubMed: 11757871]
- Sandhofer CM, Doumas LAA. Order of presentation effects in learning color categories. *Journal of Cognition and Development*. 2008; 9(2):194–221.
- Schyns PG, Goldstone RL, Thibaut J. The development of features in object concepts. *Behavioral and Brain Sciences*. 1998; 21(1):1–54. [PubMed: 10097010]
- Siskind JM. A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*. 1996; 61:31–91.
- Smith LB, Samuelson L. An attentional learning account of the shape bias: Reply to Cimpian and Markman (2005) and Booth, Waxman, and Huang (2005). *Developmental Psychology*. 2006; 42(6):1339–1343. [PubMed: 17087565]
- Stokes SF, Klee T. Factors that influence vocabulary development in two-year-old children. *Journal of Child Psychology and Psychiatry*. 2009; 50(4):498–505. [PubMed: 19017366]
- Vlach HA, Sandhofer CM, Kornell N. The spacing effect in children's memory and category induction. *Cognition*. 2008; 109(1):163–167. [PubMed: 18835602]
- Xu F, Tenenbaum J. Word learning as Bayesian inference. *Psychological Review*. 2007a; 114(No 2):245–272. [PubMed: 17500627]

Xu F, Tenenbaum J. Sensitivity to sampling in Bayesian word learning. *Developmental Science*. 2007b; 10:288–297. [PubMed: 17444970]

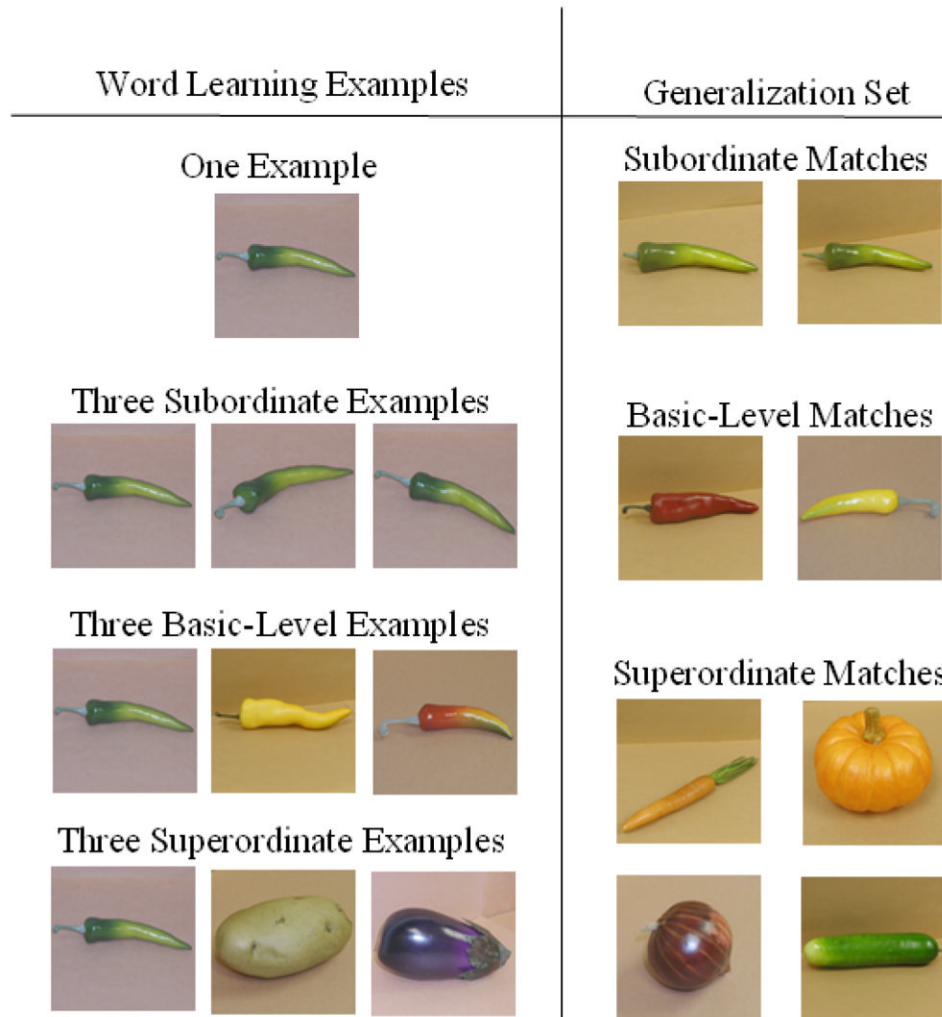


Figure 1. Exemplar set (left) and generalization set (right) for one stimulus domain (vegetables). The exemplar set is broken into four types with examples at different hierarchical levels. The generalization set was also broken into levels with different numbers of matches at each level.



Figure 2.

Examples showing the display when a single exemplar was presented (left) and when three subordinate exemplars were presented (right). Exemplars were presented in the lower portion of the screen. The generalization set was presented in a 4×6 grid in the upper portion of the screen. Participants clicked the yellow 'done' button after they had selected the matches on each trial. The selected items were highlighted by a green box (not shown).

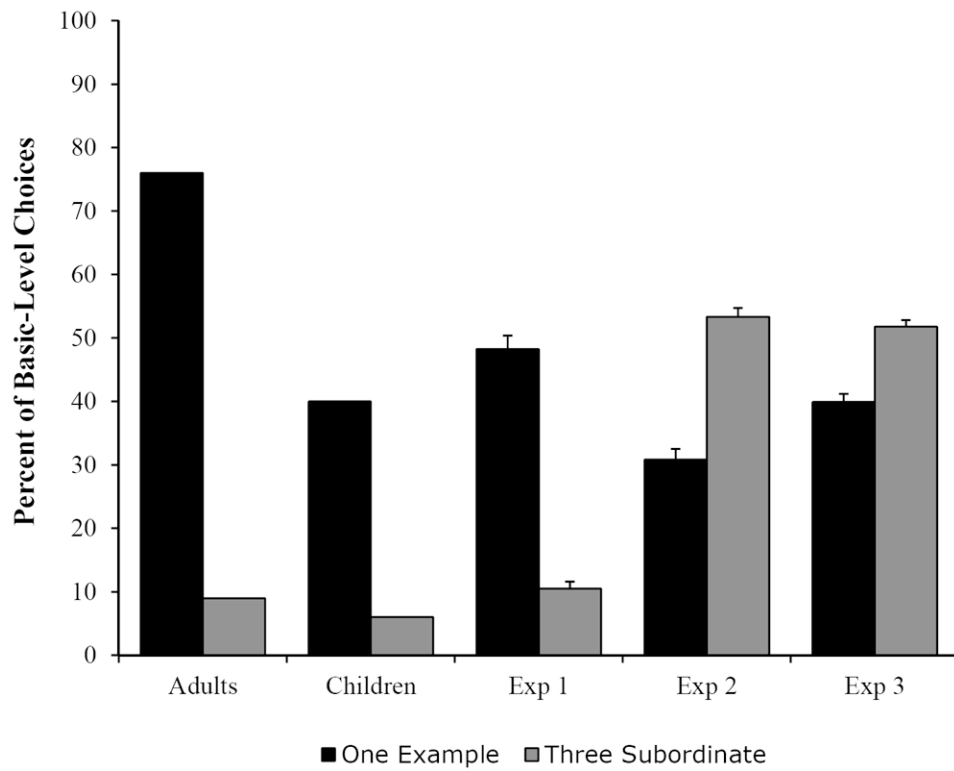


Figure 3. Percent of basic-level choices selected from the generalization set across experiments on trials with one example (black bars) and three subordinate examples (grey bars). Adult and Child data are from Xu and Tenenbaum (2007a). Error bars show standard error.

Table 1

Complete data set from Experiments 1-3, S1, and S2 showing mean percent of generalization choices at the subordinate, basic, and superordinate levels for the four types of experimental trials: when one example was presented, when three subordinate examples were presented, when three basic-level examples were presented, and when three superordinate examples were presented. The top two lines in each row show data for adults and children taken from Xu and Tenenbaum (2007a).

Generalization Choice	Experiment	Examples Presented							
		One Example	Three Subordinate	Three Basic-Level	Three Superordinate				
Subordinate	Adults	96	-	95	-	76	-	94	-
	Children	96	-	94	-	75	-	94	-
	Experiment 1	99.12	(3.82)	98.25	(5.25)	96.49	(8.92)	94.74	(16.71)
	Experiment 2	95.00	(13.35)	88.33	(16.31)	94.17	(22.47)	90.83	(26.19)
	Experiment 3	94.74	(13.66)	92.98	(16.02)	98.25	(7.64)	93.86	(13.84)
	Experiment S1	93.86	(9.95)	96.49	(11.88)	99.12	(3.82)	96.49	(8.92)
Basic	Experiment S2	91.23	(23.81)	93.86	(13.84)	91.23	(23.81)	88.60	(26.67)
	Adults					91	-	97	-
	Children					75	-	88	-
	Experiment 1					92.10	(20.31)	85.09	(27.72)
	Experiment 2	See Figure 2		See Figure 2		90.00	(13.68)	86.67	(26.27)
	Experiment 3					92.10	(15.08)	72.81	(33.43)
Superordinate	Experiment S1					82.45	(25.74)	69.30	(45.22)
	Experiment S2					85.96	(20.23)	80.70	(37.37)
	Adults	9	-	1.00	-	4	-	87	-
	Children	17	-	0.00	-	8	-	62	-
	Experiment 1	7.02	(16.01)	0.88	(2.62)	15.35	(11.2)	81.31	(23.54)
	Experiment 2	5.83	(20.42)	2.50	(4.75)	13.33	(21.01)	75.41	(30.04)
Experiment 3	4.82	(14.24)	14.47	(21.66)	17.54	(21.31)	67.98	(30.83)	
Experiment S1	6.14	(17.75)	0.88	(2.62)	21.49	(29.3)	61.40	(43.76)	
Experiment S2	1.75	(5.94)	0.00	(0.00)	11.84	(12.20)	67.26	(32.43)	