Cognitive Psychology 114 (2019) 101226

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Spotting Dalmatians: Children's ability to discover subordinate-level word meanings cross-situationally

Felix Hao Wang[a,*], John C. Trueswell[b]

[a] University of Nevada, Las Vegas, United States
[b] University of Pennsylvania, United States

ARTICLE INFO

Keywords:
Word learning
Cross-situational learning
Subordinate level meaning

ABSTRACT

Even when children encounter a novel word in the situation of a clear and unique referent, they are nevertheless faced with the problem of semantic uncertainty: when "puziv" refers to a co-present spotted dog, does the word mean *Fido, Dalmatian, dog, animal,* or *entity*? Here we explored the extent to which children (3–5 years of age) can reason about a novel word's meaning from information they have gathered *cross-situationally*, from a series of simple ostensive labeling events ("I see a puziv!"). Of particular interest were the conditions under which children arrive at a subordinate level meaning (e.g., *Dalmatian*) rather than a basic level meaning (e.g., *dog*). Experiment 1 showed that children (N = 32) were capable of using lexical contrast and/or mutual exclusivity cross-situationally, such that they arrived at subordinate level meanings only when the words being learned contrasted at the subordinate level, otherwise they strongly preferred basic level meanings (e.g., dog) even when the word had previously referred to subordinate level exemplars (always Dalmatians). Experiment 2 showed that some children in this same age range (N = 20) can also arrive at subordinate level meanings cross-situationally when offered relatively minimal linguistic support ("It's a kind of dog."). The findings are interpreted with respect to current theories of cross-situational word learning, and suggest that word meanings rather than sets of referential exemplars are tracked and used for cross-situational comparison.

1. Introduction

1.1. Semantic uncertainty in word learning

A central problem for a word learner of any age is the problem of semantic uncertainty (Quine, 1960). For example, when a speaker refers to a co-present spotted dog by saying "puziv", it is unclear whether that word means *Fido, Dalmatian, dog, animal,* or *entity*, or any of the other plausible semantic categories for which this object could serve as an exemplar, such as *pet*, *female dog, spotted animal*, *mother*, etc. The experimental data suggest that word learners, particularly children, initially gravitate toward the so-called "basic level", thinking that a novel word like "puziv", when uttered in an ostensive labeling context (i.e., a context where learners have access only to the word and its referent, e.g., "Look! A puziv!", in the presence of a Dalmatian) likely has a basic level meaning (*dog*) and not a subordinate (*Dalmatian*) or superordinate (*animal*) meaning (e.g., Golinkoff, Shuff-Bailey, Olguin, & Ruan,

* Corresponding authors at: Department of Psychology, University of Nevada, Las Vegas, 4505 S. Maryland Pkwy., Las Vegas, NV 89154-5030, United States.
E-mail addresses: hao.wang@unlv.edu (F.H. Wang), trueswel@psych.upenn.edu (J.C. Trueswell).

https://doi.org/10.1016/j.cogpsych.2019.101226
Received 2 January 2019; Received in revised form 19 June 2019; Accepted 19 June 2019
Available online 13 July 2019
0010-0285/

1995; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). So if children have a strong basic level bias when learning words from immediate observation of the referent world, how do they come to learn other concrete terms, such as subordinates like "Dalmatian" or "Collie"? Much of the existing word-learning research suggests that the learning of non-basic-level terms require special referential circumstances in which the child's interlocutor offers additional evidence regarding the word's meaning, often in the form of explicit semantic comparison and contrast. For example, observational studies reveal that adults often introduce non-basic-level terms to children by "anchoring" them to a familiar basic level term ("See this dog? It's a poodle."), sometimes explicitly mentioning the category relation (e.g., "It's a kind of dog.") (e.g., Blewitt, 1983; Callanan, 1985; Clark & Wong, 2002; Shipley, Kuhn, & Madden, 1983). This work also shows that talk of non-basic terms arises most often when multiple exemplars of the same category are encountered together in the referent world, such that parents are more likely to use superordinate terms to label groups ("See the animals?"), offering the opportunity to identify perceptual and semantic commonalities among members of the set. Moreover, reference to members of the group are often distinguished perceptually or semantically ("And this is a Dalmatian, see the spots?") (see especially, Callanan, 1985; Shipley et al., 1983).

If these parental production patterns reflect the conditions under which subordinate terms are learned, we would expect language novices to assume that a novel word has a subordinate meaning only when some or all of these conditions are met. Indeed, this appears to be the case. For example, when Waxman, Shipley, and Shepperson (1991) taught children (3-yr-olds) the meanings of novel nouns in simple ostensive labeling contexts (Exp 1., e.g., naming a series of poodles each with the same novel noun), they showed a strong basic level bias, thinking for instance that the novel words meant *dog*, not *poodle*. However, children correctly learned subordinate meanings in two conditions: either when sibling subordinate terms were contrasted, e.g., by highlighting the known differences of breeds of dogs while labeling them with two different nonsense terms (Exp. 2, e.g., "Daxes help take care of sheep. Mipens find birds for hunters") or by anchoring novel terms to the known basic level terms ("This is a dax. It's a kind of dog."; Exp. 3). In all three experiments, relevant referential exemplars were always co-present during labeling (e.g., a Dalmatian and poodle side by side), presumably to help children with perceptual comparison between referential exemplars (see also Waxman, 1990, for related work, in which use of novel nouns alone bias children away from subordinate meanings). Callanan (1989) reported a word learning study with very similar results to Waxman et al. (1991). In that study, children (3-year-olds) learned the meaning of novel labels from groups of co-present referential exemplars. Simple ostensive labeling led to basic level meanings, whereas talk of semantic inclusion (i.e., use of the linguistic anchor "kind of") led children to non-basic level meanings (superordinate meanings, in this case).

Notably, other work shows that younger children appear to have a weaker understanding of "kind of", such that 2-year-olds respond to inclusion statements ("A dax is a kind of mipen") as mutually exclusive terms (Diesendruck & Shatz, 2001, see also Gottfried & Tonks, 1996). Thus, talking about semantic inclusion involving co-present exemplars allows children to zoom in on the target semantic category, at least when directed to a child who is equipped with the linguistic ability to correctly interpret these expressions (knowing what "kind of" means, and/or knowing what "dog" means). Gelman, Wilcox, and Clark (1989) made similar observations, but focused on the role of contrast and exclusivity at a given semantic level as being especially important. They taught children subordinate meanings by linguistic support: "(Pointing to a dingo) This is a dog, it's a dingo; (now pointing to a non-dingo dog) This is a dog, it's not a dingo." They found that children sometimes assign the two words to two separate meanings (dingo meaning only dingo and dog meaning non-dingo dogs), rather than assigning the two words to two meanings at different levels. They argued that this behavior was in line with word learning principles such as "contrast" (Clark, 1987, 1988) and "mutual exclusivity" (Markman, 1987). This kind of meaning assignment was argued to have arisen from children's reluctance to conclude that two different words may have the same meaning or same extension. But note that children nevertheless arrived at a non-basic category level given the presence of linguistic support.

To summarize, the work reviewed here suggests that the learning of subordinate terms occurs under more specific referential circumstances that contain comparison and contrast – often under situations in which an interlocutor explicitly mentions this information in the presence of groups of referents. This suggests that the ability to learn which words label non-ordinate categories will only be accomplished by a more experienced language learner who has the sophistication to interpret contrastive expressions and/or predicates that convey subset relations (e.g., "That's a Dalmatian. It's a kind of dog."), and indeed the literature suggests that this is the case. In the general discussion, we return to the issues of learning from observation of the referent world to discuss in more detail the source of the so-called basic level bias during early word learning, but suffice it to say, when contrastive information is absent from the situation, it is expected that any learner, no matter their linguistic and conceptual sophistication, will gravitate toward basic level meanings for novel words. In particular, simple ostensive labeling events (such as seeing a Dalmatian and hearing "That's a puziv.") are expected to lead learners to assign a basic level meaning for that word (*dog*).

1.2. Learning subordinate meanings cross-situationally: a Bayesian account

Note that the above work emphasizes children's reliance on information from the immediate situational context in which the word is being heard. Children work with the local context to guess the word's meaning, and presumably this meaning (i.e., the decided upon semantic category) is what is stored with the word and later retrieved on the word's next occurrence. However, more recent studies have proposed an alternative way for arriving at the accurate meaning of a word, in which information is gradually accrued from multiple encounters with a word, and statistically compared "cross-situationally" (e.g., Siskind, 1996; Yu & Smith, 2007). An influential paper by Xu and Tenenbaum (2007) explored the possibility that such a learning procedure could aid in identifying subordinate level meanings, even from a series of simple ostensive labeling events. They proposed that non-basic level terms, such as subordinate terms, could be acquired from situations of ostensive labeling alone, as long as one assumes the learner tracks the referential exemplars of a word over multiple labeling occurrences and makes Bayesian-like inferences about how these

exemplar patterns do or do not support category memberships. They offered the following hypothesis:

> "*We argue that this essential problem … can be solved by a Bayesian approach to word learning… To illustrate with the ostensive learning problem … after observing Max the Dalmatian labeled as a fep, a learner guided by a taxonomic hypothesis space and perhaps some preference for labeling basic level categories might reasonably guess that fep refers to all dogs. Now suppose that the learner observes three more objects labeled as feps, each of which is also a Dalmatian. … after these additional examples, the word fep seems relatively more likely to refer to just Dalmatians than to all dogs. 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 & Tenenbaum, 2007)

Under this account, the learner stores the referential exemplars that they encounter with any particular word, so that they can later evaluate the likelihood that those exemplars come from different categories. In the example above, all the exemplars for "fep" were Dalmatians. According to this "suspicious coincidence" account, "fep" is thus more likely to mean *Dalmatian* than *dog* because the likelihood of sampling three Dalmatians from the category *dog* is lower than sampling three Dalmatians from the category *Dalmatian* (i.e., it would be a "suspicious coincidence" to have sampled three dogs from the exact same breed). As they explained, such an inference follows naturally from Bayesian statistical inferencing about hypothesized meanings cross-situationally.

Xu and Tenenbaum (2007) offered compelling evidence in favor of this account, showing that after encountering several Dalmatians all labeled as "feps", children (3 and 4 years of age) had a strong preference for the subordinate meaning: when asked to find more "feps" they selected Dalmatians 96% of the time and other dogs only 4% of the time. Although this is a compelling result, it is worth noting that it was not obtained from a study of cross-situational word learning of the sort one might expect, in which individual Dalmatians were encountered and labeled as a "fep" (e.g., on three or four different occasions), which would be the central idea of what it means to learn cross-situationally. Instead, critical trials involved reference to a group of three co-present Dalmatians, all labeled as "feps". This was done by having the experimenter first select Dalmatians from a group of other toy animals, some of which were dogs which were also co-present with the utterance. After being labeling as "feps", and with the other toys still present, children were asked to find "more feps". Thus, although the work invites an interesting conclusion about Bayesian inferencing cross-situationally, the test procedure did not align as well as it could with the claim, and, in fact, was conducted in an immediate situational context that contained many of the specialized circumstances that past work found important for identifying subordinate meanings (i.e., groups of co-present referents, talked about in a way that encourages comparison within and across groups). Notably, although there have been successful replications of the Xu and Tenenbaum (2007) study, all of these other experiments (Jenkins, Samuelson, Smith, & Spencer, 2015; Lewis & Frank, 2016) were conducted under similarly rich contrastive settings and not done via cross-situational word learning in which individual referents were encountered and labeled. Moreover, work with adults suggests that presenting exemplars simultaneously as a group helps to encourage subordinate meanings (Spencer, Perone, Smith, & Samuelson, 2011; see also Christie & Gentner, 2010, for a similar phenomenon involving structural alignment in children), further suggesting a need for perceptual comparison of co-present referents.[1]

Although the current literature does not allow us to conclude whether or not children are able to learn subordinate meanings cross-situationally via Bayesian inference, it would be very important to know if young children could indeed make use of such a procedure to learn subordinate meanings. If this mechanism is available for young word learners, it would mean that they would have the means to learn subordinate meanings on their own without having to wait until they find the specific referential circumstances to learn subordinate meanings. In other words, children would not need to wait until their interlocutors (expert language users) offered additional linguistic or contextual contrastive evidence regarding the subordinate meanings of words. However, there is reason to believe that such a mechanism of cross-situational comparison might not be available for young children. First, limitations in young children's working memory and long-term memory (e.g., Gathercole, Pickering, Ambridge, & Wearing, 2004) might lead one to expect that children in this age range would have difficulty accomplishing cross-situational comparison in which past referential exemplars for a word are retrieved and compared to determine the plausible semantic category for a word.

Second, recent work on cross-situational comparison suggests that, at least when it comes to reducing *referential* uncertainty (rather than semantic uncertainty), children, 2 and 3 years of age, exhibit highly limited and focused memory abilities, recalling only the previously selected referent from a past encounter with a word, and not referential alternatives that were also present during that encounter (Woodard, Gleitman, & Trueswell, 2016). It is possible that similarly limited and focused memory abilities would be present in children trying to recall semantic alternatives of previously encountered referents (even ones they selected), making it less likely that they could detect the suspicious coincidence that, e.g., all past referents were not only dogs, but also Dalmatians, and apply this to word learning.

[1] It is notable that in another study of adults, Lewis and Frank (2016) found that 1-sec sequential presentation of multiple subordinate exemplars does not completely eliminate the suspicious coincidence effect (contra the stronger conclusions of Spencer et al., 2011). Lewis and Frank nevertheless found an effect of simultaneous vs. sequential presentation of referential exemplars in the first block of trials in all experiments, such that simultaneous presentation of multiple referents increases the preference for subordinate meanings (see Fig. 2, Lewis and Frank, plus further analyses reported in Caplan, submitted). Moreover, like all other studies of the suspicious coincidence effect, exemplars in the Lewis and Frank study were presented in the presence of other potentially contrasting referents (i.e., the co-present test array). Thus, contextual factors that are known to support subordinate level word meanings (e.g., Callanan, 1985; Waxman et al., 1991) were present in this work as well.

1.3. Testing how subordinate meanings might be learned cross-situationally

Although the Xu and Tenenbaum (2007) Bayesian inference account was intended as a theory of cross-situational word learning, work to date has not tested whether inferences of this sort can actually be accomplished cross-situationally. We set out to conduct such studies here, with children ages 3–5 years. This age was selected because it is the same age tested in Xu and Tenenbaum (2007) and is the very age at which children show signs of learning subordinate meanings, at least when helpful contrastive information is presented during the learning instance itself, i.e., not cross-situationally (e.g., Waxman et al., 1991). The present work was designed to establish the extent to which the learning of subordinate level meanings can be accomplished from sampling referential exemplars cross-situationally from a series of simple ostensive labeling events ("I see a dax!"). As will be presented below, we used the simplest labeling event: each encounter with a novel word was done in an environment in which there was a single, clear and unique referent. And, we tried the simplest test of the prediction: comparing a sequence in which three Subordinate Exemplars are encountered (three Dalmatians) to a sequence in which a variety of exemplars are encountered (three dogs from different breeds). If children engage in reasoning about the sampling statistics of referential exemplars in the way proposed by the suspicious coincidence account, we should expect children to select the subordinate level (*Dalmatian*) meaning from the former condition and the basic level (*dog*) meaning from the latter condition.

Our experiments further test the robustness of cross-situational comparison by examining whether the more traditional proposals for how children learn subordinate terms (i.e., contrast and linguistic support) also apply to circumstances in which children are required to constrain word interpretations cross-situationally, rather than in the immediate co-presence of pairs of referents. In particular, the work of Gelman et al. (1989) indicates that the learning of two novel words that contrast at the subordinate level can result in children arriving at mutually exclusive subordinate meanings for both words. However, these and other positive findings of using contrastive information (e.g., Waxman et al., 1991) were observed in settings in which the referential exemplars were co-present (e.g., seeing two objects and hearing "This is a dax, and this is a mipen."). It is important to know if children can also show such abilities when word-referent pairs are encountered on separate labeling events, at least in a context in which the two words of interest are uttered in close temporal proximity. If so, it would suggest that detecting the special circumstances that support subordinate meanings is robust enough to memory demands associated with encountering two separate labeling situations within the same dialogue.[2] Note that detecting lexical contrast only requires children to recall what another novel term means and to notice that it is the same as the probable meaning of the present novel word. Such a procedure when applied cross-situationally does introduce memory demands because they have to recall the meaning of another, recently uttered novel word. But these memory demands are arguably substantially less than what would be needed for detecting suspicious coincidences cross-situationally, i.e., retrieving sets of referential exemplars and drawing an inference about this set. Thus, in Experiment 1 below, we also examine if children can detect and use lexical contrast cross-situationally from a series of separate labeling events, asking children to learn words that contrast either at the basic level or at a subordinate level.

After exploring these issues in Experiment 1, Experiment 2 takes up the issue of using linguistic support to identify subordinate word meanings cross-situationally. Past work has found that linguistic support is helpful. In particular, anchoring novel terms to familiar basic level terms ("This is a dax. It's a kind of dog.") helps children compare word meanings locally and helps them arrive at subordinate meanings for novel words, especially in older children (e.g., Waxman et al., 1991). But again, this finding was obtained in complex referential contexts of co-present referents in which the referents were contrasted semantically. Experiment 2 therefore examines the role of using inclusive terms in a cross-situational word-learning task, in which each labeling event simply refers to a single co-present referent.

2. Experiment 1

Experiment 1 explored whether children could identify subordinate level meanings cross-situationally from a series of simple ostensive labeling events. We tested thirty-two 3-year-old to 5-year-old children (mean age 49.3 months), which was the same age range as that used in the child study of Xu and Tenenbaum (2007) and is an age during which children are showing rapid vocabulary growth (e.g., Hart & Risley, 1995). In the present study, children played a game on a computer in which they were to help two puppets, Aaron and Sally, put pictures in Sally's scrapbook. On each exposure trial, the child saw an image of an object or an animal on the screen and Sally labeled it with a novel word, saying, e.g., "Oh wow! That's a puziv! I see a puziv! Do you see a puziv?" (see Fig. 1A below). Six exposure trials (three for each word) occurred, such that the child never saw multiple exemplars at the same time (just one exemplar per trial), alternating between the two novel words. After this, the test phase began. A set of eight images was presented above Sally's scrapbook (Fig. 1B) and Sally instructed the child to put images in her scrapbook using the novel words. Among the eight images were two subordinate level exemplars (e.g., two Dalmatians) and two basic level exemplars (e.g., two non-Dalmatian dogs). Selection of the Dalmatians but not the other dogs would be consistent with a Subordinate level meaning. Selection of the Dalmatians and the other dogs would be consistent with a Basic level meaning. Selection of the other dogs but not the Dalmatians would be consistent with a meaning that is Mutually Exclusive with the subordinate meaning.

[2] This essentially requires children to understand that within a local conversational context, words likely contrast at a particular level within a semantic hierarchy (daisies as opposed to roses). This claim finds support in elegant studies by Waxman and Hatch (1992) who find children (3–4 years) do exactly this in their own productions of known words. It remains an open question whether they can apply such conversational logic to cross-situational word learning.

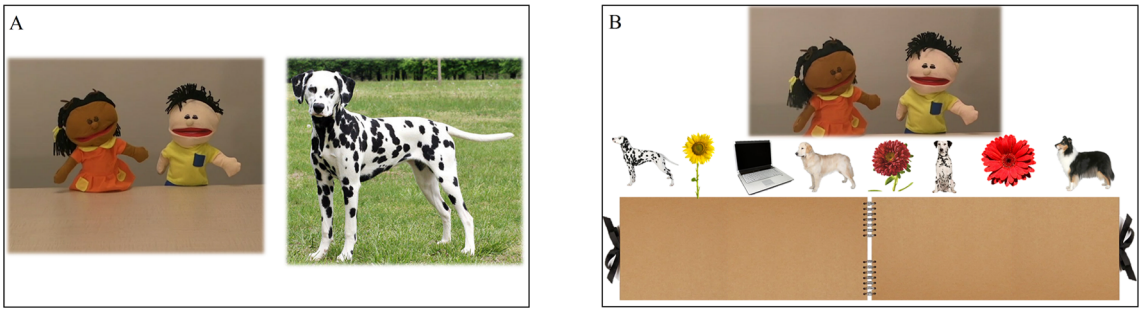

Fig. 1. Screen shots of the learning trials from the exposure phase (A) and test phase (B).

There were two conditions in Experiment 1, which differed in what was encountered during the exposure phase (see Fig. 2). For all children regardless of condition, one of the two novel words was always paired with exemplars of a subordinate category (e.g., three different Dalmatians, each labeled as a "puziv"), whereas the second novel word was always paired with exemplars of a basic level category. For children in the Contrast at the Basic Level condition, the second word labeled exemplars from a different semantic domain than the target word (e.g., three different birds, each labeled as a "ruften"). Thus, in this condition, the two novel words plausibly contrasted at the Basic level (e.g., "puzivs" were dogs whereas "ruftens" were birds). However, note that the first word could just as well belong to the Subordinate Level Category (e.g., "puziv" could mean *Dalmatian*, since after all, this word was paired with a Dalmatian on all three occurrences of that word (a situation very close to the experiment of Xu and Tenenbaum (2007), with the exception that in their experiment all three Dalmatians were presented together, while in our study each Dalmatian appeared individually, one after the other). In the Contrast at the Subordinate Level condition (Fig. 2B) however, the second novel word was paired with Basic Level exemplars from the same semantic domain as the other novel word (e.g., three different dogs, each a different (non-Dalmatian) breed). As such, the two words might plausibly contrast at the Subordinate level (e.g., "puziv" could mean *Dalmatian* whereas "ruften" could mean *other dogs*).

Of interest are the meaning selections children make at test. We offer detailed predictions below, but, in brief, if children do not engage in cross-situational comparison, they should think all words have a basic level meaning at test, regardless of the exemplar types encountered or the contrastive context. This is because each ostensive labeling context is consistent with the basic level meaning and hence it is simply reinforced on each additional learning instance. If, on the other hand, children can compare the meanings of different novel words across each occurrence (a minimal form of cross-situational comparison), and if they prefer that novel words have different meanings (in accordance with *contrast* or *mutual exclusivity*), then they should have a preference for basic level meanings in the Contrast at the Basic Level condition, but a preference for Subordinate Level, mutually exclusive, meanings in the Contrast at the Subordinate Level condition. This is because in the Contrast at the Basic Level condition (Fig. 2A), the ostensive labeling supports two different basic level meanings ("puziv" = *dog*, "ruften" = *bird*) whereas in the Contrast at the Subordinate Level condition (Fig. 2B), these individual labeling events on their own support the same Basic Level meaning ("puziv" = *dog* and "ruften" = *dog*) which should lead children to search for two non-overlapping meanings for these words, in particular "puziv" should have a Subordinate meaning (*Dalmatian*) whereas "ruften" should have a Mutually Exclusive meaning (*non-Dalmatian dog*).

Finally, if children track the referential exemplars encountered with each word and use Bayesian inference to identify probable meanings, then children should think a word paired with 3 Subordinate Level exemplars has a Subordinate Level meaning whereas a

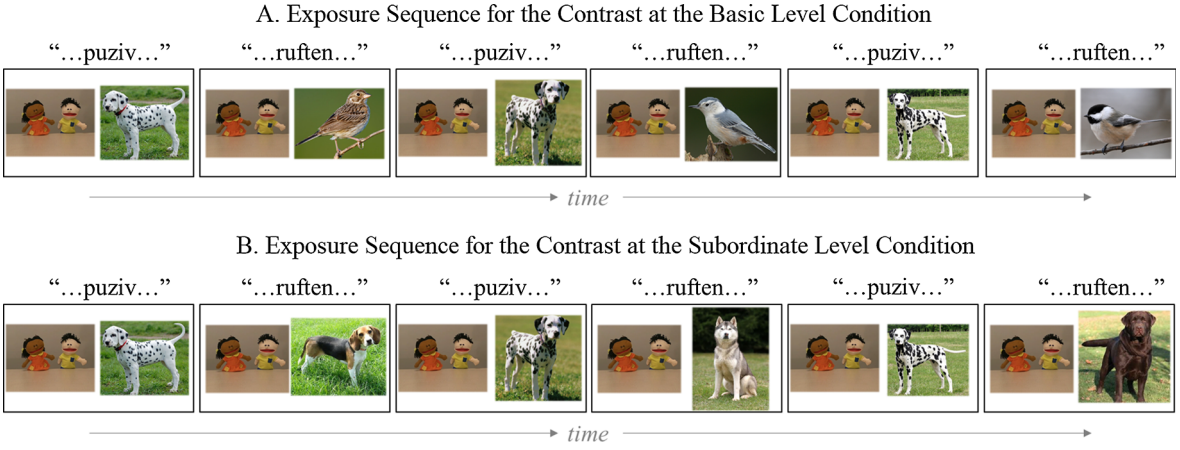

Fig. 2. Screen shots of the trial sequence during the Exposure Phase, for the Contrast at the Basic Level condition (A) and the Contrast at the Subordinate Level condition (B).

word paired with 3 Basic Level exemplars has a Basic Level meaning. For example, in the Contrast at the Basic Level condition (Fig. 2A), it would be a suspicious coincidence if three Dalmatians were a random sample of dogs, and hence "puziv" should mean *Dalmatian*, not *dog.* And, it would *not* be a suspicious coincidence if three different kinds of birds were a random sample of birds, and hence "ruften" should mean *bird.* Note that the same general expectation holds for the Subordinate Contrast Condition (Fig. 2B). It would again be a suspicious coincidence if three Dalmatians were a random sample of dogs, and hence "puziv" should mean *Dalmatian.* And, it would *not* be a suspicious coincidence if three different kinds of dogs were a random sample of dogs, and hence "ruften" should mean *dog.* Note that it is not clear if the presence of potentially contrastive word meanings should modulate this predicted pattern, as the account that was developed by Xu and Tenenbaum (2007) did not consider such a possibility. Thus, an effect of Contrastive context cannot be seen as refuting, nor supporting, such a theory. What is clear though is that any plausible version of this theory predicts children in this experiment should have a Subordinate Meaning preference for words paired with 3 Subordinate Exemplars, including those children assigned to the Contrast at the Basic Level condition. This is because, in this condition, it is not apparent why learning one word means *bird* would dissuade a child from proposing that another word means *Dalmatian*, given the very suspicious coincidence that every single time the word had been heard, it was paired with a Dalmatian.

2.1. Method

2.1.1. Participants

Thirty-two typically developing children between the age of 3 and 5 years were recruited from preschools in the Philadelphia area (mean age = 49.3 months; Standard Deviation = 8.2 months, range = 36–60 months). All were native speakers of English. Children were assigned to one of the two experimental conditions, i.e., Contrast at the Basic Level (N = 16, mean age 48.7) or Contrast at the Subordinate Level (N = 16, mean age 49.9). Assignment was random, except so as to match the two groups based on age and gender. Parents provided informed consent in accordance with an IRB protocol of the University of Pennsylvania. Children were given a small toy as compensation for their participation.

2.1.2. Stimuli

In order to create visual stimuli for the exposure phase of the experiment, we selected photographic images from the internet that depicted exemplars of six pairs of contrasting categories (i.e., six Subordinate Level categories paired with six Basic Level categories): *Dalmatian* vs. other types of *dog*, *police car* vs. other types of *car*, *orange Persian cat* vs. other types of *cat*, *goldfish* vs. other types of *fish*, *basketball* vs. other types of *ball*, and *parrot* vs. other types of *bird*. For each of the resulting 12 categories, we selected three photos, each depicting a single exemplar centered in the photo and appearing within a naturally occurring background. The three exemplars for each Subordinate Level category had to be easily recognizable members of that category. The selection of three images for each Basic Level category had additional requirements. Each had to be a member of a different subordinate level category of the Basic Level category, and each could not be a member of our experimentally designated Subordinate Level category. E.g., for dogs, we selected exemplars of three different dog breeds, none of which were Dalmatians; and for birds, exemplars of three different bird species, none of which were parrots. For the test phase of the experiment, we selected two additional photos per category, each depicting a single exemplar that was not presented before (i.e., during the exposure phase). These test images were selected in the same way as the exposure images, except each exemplar was tightly cropped (no background), so that they could be dragged on the screen as objects during test.

A female native speaker of Standard American English recorded the auditory stimuli. Each exposure phase image was paired with a recording of the speaker using a nonsense noun three times: "Oh wow! That's a ruften! I see a ruften. Do you see a ruften?" Test utterances, to be used in the test phase, consisted of the same speaker saying "I want the ruftens. Can you put the ruftens in the scrapbook?" The utterances were recorded in a sound attenuating room and digitized at a sampling rate of 44.1 kHz. Separate recordings were made for six disyllabic non-words (*antow, channer, lartsue, mirshow, puziv, ruften*).

2.1.3. Procedure

The experimental procedure consisted of a warm up phase, followed by three rounds of exposure and test.

Warm up phase. Each child was tested individually in a quiet room either at a local school or in the lab. The child was seated in front of a computer screen, with the experimenter seated next to the child. The child was told they were going to play a game on the computer in which they would help two puppets, Aaron and Sally, put pictures in Sally's scrapbook. The child then engaged in a "warm up" phase designed to familiarize them with the test procedure, using known words. In particular, the child watched a video of two puppets, Aaron and Sally, who had an empty scrapbook that appeared below them. Aaron (addressing the child) explained that he wanted to put pictures in Sally's scrapbook, but needed help to do this. In the first warm-up trial, eight photos appeared above the scrapbook, consisting of three cookies, three bicycles and two doors (arranged in a way similar to Fig. 1B above). Aaron then said "I want the bikes. Can you put the bikes in the scrapbook?" The experimenter then asked the child to touch the bikes on the screen, which resulted in each bike moving down to the scrapbook, via simple animation on PowerPoint. If the child didn't select all the bicycles, the experimenter asked, "Are there any more bicycles?", which then resulted in selection of the rest of the bicycles. After positive feedback from the experimenter, the task repeated with Aaron saying "I want the cookies. Can you put the cookies in the scrapbook?" All children performed this task correctly.

Exposure phase. Immediately after the warm up, the experiment transitioned to the exposure phase, where Aaron announced that now Sally, rather than he, wanted to put pictures in the scrapbook. To motivate the use of novel words, Aaron explained that Sally did not speak English very well and instead speaks puppet talk. Aaron proposed that Sally should teach everyone some puppet talk so that

they could help put pictures in the scrapbook. Children were asked if they would like to help Sally with the task, and the exposure phase began after they said yes.

The exposure phase consisted of six trials in which children attempted to learn two novel words (e.g., "puziv" and "ruften"), three trials per word, six trials in total. On each exposure trial, a single exposure photo (e.g., of a Dalmatian) appeared to the right of the puppets (Fig. 1A). Sally would then look over at the photo and label it three times with a novel word (e.g., "Oh wow! That's a puziv! I see a puziv! Do you see a puziv?"). After the utterance ended, there were two seconds of silence during which the child could examine the exemplar. Then, the next exposure trial would occur, involving the second word (e.g., "Oh wow! That's a ruften! I see a ruften! Do you see a ruften?"). Exposure trials would alternate between the two words to be learned, resulting in six trials in total (see Fig. 2). For all children regardless of condition, one of the two novel words was always paired with exemplars of a subordinate category (e.g., three different Dalmatians, each labeled as a "puziv") whereas the second novel word was always paired with exemplars of a basic level category. For children in the Contrast at the Basic Level condition (Fig. 2A), the second word labeled exemplars from a different semantic domain than the target word (e.g., three different birds, each labeled as a "ruften"). In the Contrast at the Subordinate Level condition (Fig. 2B), this second word instead labeled basic level exemplars from the same semantic domain (e.g., three different dogs, each a different (non-Dalmatian) breed).

Test phase. Immediately after the exposure phase, the test phase began. To assess the meanings that children had assigned to the novel words during the exposure phase, the puppets asked the child to help them put pictures in Sally's scrapbook (Fig. 1B). Each word was tested separately. For instance, at the start of a test trial, children would hear Sally say "I want the puzivs. Can you put all the puzivs in the scrapbook?", note that in the test all the novel words were presented in the plural form (e.g., "puzivs", "ruftens"), while in the exposure phase each novel word was presented in its singular form because each picture only showed one possible referent. The array of test images (8 in total) appeared below Sally and included two Subordinate Level test images (Subordinate Level matches), two Basic Level test images (Basic Level matches), three distractor test images from another basic level), and another distractor test image from another basic level. Children were allowed to keep choosing the test images that they wanted to put on the scrapbook until they were done.

Repetition of Exposure and Test. Two additional blocks of Exposure and Test were administered with new words and new exemplars, resulting in three blocks in total. Like the first block, each additional block involved the learning of two novel words, one at the subordinate level and one at the basic level. Thus, by the end of the experiment, children had been tested on six different novel words, three having Subordinate Level exemplars and three having Basic Level exemplars.

2.1.4. Design

For the Contrast at the Basic Level condition, two experimental lists were constructed, which swapped the semantic domain used for each word. For example, in List 1, "puziv" was paired with Dalmatians (Subordinate exemplars) and "ruften" with birds (Basic Level exemplars), whereas in List 2, "puziv" was now paired with non-Dalmatian dogs (Basic Level exemplars), and "ruften" with Parrots (Subordinate exemplars). Children were assigned to one of the two lists.

Because children who were assigned the Contrast condition were exposed to two words from the same semantic domain in each block (e.g., two words referring to dogs) and because there were only three blocks, children in this condition were exposed to only three of the six semantic domains. Thus, in the Contrast condition, two experimental lists were created, such that one list involved three of the semantic domains (dogs, cars and fish) and the other list involved the other three domains (birds, balls and cats). Children were assigned to one of the two lists.

2.1.5. Coding of test responses

On test trials, children were allowed to select any number of the eight pictures presented in the test array (see Fig. 1B). These pictures included two Subordinate Level matches (e.g., two Dalmatians), two Basic Level matches (e.g., two non-Dalmatian dogs) and four other pictures that depicted 'incorrect' referents of two other basic level categories that were not mentioned during the exposure phase. We assigned each child's selection pattern on a given trial to one of three theoretically relevant categories, which we describe here with respect to the example given in Fig. 1B. In particular, when asked to select "all the puzivs", a response was deemed to convey a Basic Level Meaning if the child selected all the dogs, i.e., both the Subordinate Level Matches (the Dalmatians) and the Basic Level Matches (the non-Dalmatian dogs). A response was deemed to convey a Subordinate Level Meaning if the child selected the Dalmatians but none of the non-Dalmatian dogs, i.e., the Subordinate Level Matches but none of the Basic Level Matches. A response was deemed to convey a Mutually Exclusive Meaning if the child selected the non-Dalmatian dogs but none of the Dalmatians, i.e., the Basic Level Matches but not the Subordinate Level Matches. To simplify our coding, any child who missed one of the two matches within a category (e.g., selecting only one of the two Dalmatians) was coded as if he/she had selected both matches. For example, a child who selected one Dalmatian and two non-Dalmatian dogs (i.e., a child who missed one of the Dalmatians) would have been coded as conveying a Basic Level Meaning; and, e.g., a child who selected just one Dalmatian and no other dogs would have been coded as conveying a Subordinate Level Meaning. (Missing an item in such a way was rare, happening only on 5.2% of all trials.) Responses were excluded from analysis if the child selected any of the distractor images in addition to, or instead of, the target images. This procedure excluded 15% of all the data. This means that children on average correctly learned the semantic domain of 85% of all words.

2.1.6. Predictions

The predictions for the two theoretical accounts of interest appear in Table 1 below. Consider first what we can call the "Contrast Account", namely the hypothesis that children can compare the meaning of the words across different exposure trials (a minimal form

Table 1
Predicted meaning preferences at test for the two theoretical accounts of interest.

	A. Contrast at the Basic Level Condition During Exposure	
	3 Subordinate Level Exemplars "Puziv" "Puziv" "Puziv" 	3 Basic Level Exemplars "Ruftin" "Ruftin" "Ruftin" 
Predictions at test for…		
Contrast Account:	Basic level meaning ("Puziv" means *dog*)	Basic level meaning ("Ruftin" means *bird*)
Suspicious Coincidence Account:	Subordinate meaning ("Puziv" means *Dalmatian*)	Basic level meaning ("Ruftin" means *bird*)

	B. Contrast at the Subordinate Level Condition During Exposure	
	3 Subordinate Level Exemplars "Puziv" "Puziv" "Puziv" 	3 Basic Level Exemplars "Ruftin" "Ruftin" "Ruftin" 
Predictions at test for…		
Contrast Account:	Subordinate meaning ("Puziv" means *Dalmatian*)	Mutually Exclusive meaning ("Ruftin" means *non-Dalmatian-dog*)
Suspicious Coincidence Account:	Subordinate meaning* ("Puziv" means Dalmatian) *possibly very strong preference	Basic level meaning ("Ruftin" means dog)

of cross-situational comparison) and prefer words to have different meanings (i.e,. *contrast*, Clark, 1987, 1988; or mutual exclusivity, Markman, 1987). In the Contrast at the Basic Level condition (Table 1A), each individual labeling event during the exposure phase supports a basic level meaning, resulting in the two words having different basic level meanings. Thus, at test, children should be content to assign basic level meanings to the two words (e.g., “puziv” means *dog*, “ruften” means *bird*). However, in the Contrast at the Subordinate Level condition (Table 1B), each individual labeling event points to the same basic level meaning (both words mean, e.g., *dog*) leading children to search for distinctive meanings. Past, non-cross-situational work (Gelman et al., 1989; Waxman et al., 1991), suggests that children in this age range arrive at two non-overlapping meanings for word pairs in such situations, and as such the first word (e.g., “puziv”) should have a Subordinate Level Meaning (e.g., *Dalmatian*) and the second word (e.g., “ruften”) should have a Mutually Exclusive meaning (e.g., *Non-Dalmatian dog*). Such a finding would only be possible if children can achieve this reasoning cross-situationally, across distinct learning instances.

Now consider the predictions if children are using a form of Bayesian inference to identify probable word meanings cross-situationally (labeled the “Suspicious Coincidence Account” in Table 1). This theory (Xu & Tenenbaum, 2007) proposes that children store referential exemplars across word occurrences and use the set of exemplars to infer the most plausible meaning at test. In the Contrast at the Basic Level condition (Table 1A), one word was paired with 3 Subordinate Level Exemplars (e.g., 3 Dalmatians) and the other with 3 Basic Level Exemplars from a different semantic domain (e.g., 3 birds of different species). It would be a suspicious coincidence if the three Dalmatians represented a random sample of dogs, and thus this word should be assigned a Subordinate meaning at test (e.g., “puziv” means *Dalmatian*). The three exemplars for the other word are a reasonable sample of birds generally,

Table 2
Proportion of Meaning Responses at Test. Experiment 1. Dominant response in bold.

	A. Contrast at the Basic Level Condition	
	During Exposure	
	3 Subordinate Level Exemplars "Puziv" "Puziv" "Puziv" 	3 Basic Level Exemplars "Ruftin" "Ruftin" "Ruftin" 
Proportion of responses at test for…		
Subordinate Level Meaning	0.025	0.025
Basic Level Meaning	**0.975**	**0.975**
Mutually Exclusive Meaning	0.000	0.000

	B. Contrast at the Subordinate Level Condition	
	During Exposure	
	3 Subordinate Level Exemplars "Puziv" "Puziv" "Puziv" 	3 Basic Level Exemplars "Ruftin" "Ruftin" "Ruftin" 
Proportion of responses at test for…		
Subordinate Level Meaning	**0.525**	0.256
Basic Level Meaning	0.400	0.333
Mutually Exclusive Meaning	0.075	**0.410**

and as such this word should be assigned a Basic Level meaning (e.g., “ruften” means *bird)*. Note that the same logic holds for the Contrast at the Subordinate Level condition (Table 1B). In this condition, one word had been paired with 3 Subordinate Level Exemplars (e.g., 3 Dalmatians) and the other paired with 3 Basic Level Exemplars from the same semantic domain (e.g., 3 other dogs). It would again be a suspicious coincidence if the three Dalmatians represented a random sample of dogs, and thus this word should be assigned a Subordinate Meaning at test. The three exemplars for the other word, on the other hand, are a reasonable sample of dogs generally, and thus this word should be assigned the Basic level meaning *dog*. Finally, although the Xu and Tenenbaum account does not make specific predictions about differences between the two Contrast conditions (Contrast at the Basic Level vs. Contrast at the Subordinate Level), the most reasonable expectation from the above discussion is that the Contrast at the Subordinate Level condition should strengthen the preference for a Subordinate Level Meaning for “puziv” even further (i.e., for the word that was paired with 3 Subordinate Level Exemplars).

2.2. Results and discussion

2.2.1. Overall pattern of meaning responses

As explained above, children’s meaning responses at test were expected to fall into one of three types: Subordinate Level Meaning (e.g., believing the word means *Dalmatian*), Basic Level Meaning (e.g., believing the word means *dog*), or Mutually Exclusive Meaning (e.g., believing the word means *non-Dalmatian-dog*). Table 2 presents the proportion of each type of response across the two types of semantic contrast. Responses from children in the Contrast at the Basic Level condition are in the top half of the table, and responses from children in the Contrast at the Subordinate Level condition are in the bottom half. In each condition, data are presented separately for words that had been paired during the exposure phase with 3 Subordinate Level Exemplars (left) or 3 Basic Level Exemplars (right).

As can be seen in the top half of Table 2, children in the Contrast at the Basic Level condition rarely assigned Subordinate level meanings to words (2.5% of the time, row 1), and instead showed an overwhelming preference for Basic Level meanings (97.5% of the time, row 2). This occurred regardless of whether 3 Subordinate exemplars or 3 Basic level exemplars had been encountered during the exposure phase. Thus, in this condition children failed to detect the suspicious coincidence that three Subordinate Level exemplars (e.g., three Dalmatians) had been paired with the nonsense word (e.g., “puziv”). Instead, they preferred the basic level meaning (e.g., *dog*) in accordance with what would be expected from the “Contrast Account”, i.e., what would be expected if they attended to the fact that each labeling event supported a basic level meaning and the contrast supported this as well - rather than

calculating fine-grained Bayesian inferences by storing exemplars that they encountered with any particular word, to later evaluate the likelihood that those exemplars came from various categories.

As can be seen in the bottom half of Table 2, children in the Contrast at the Subordinate Level condition were much more likely to assign a Subordinate meaning to the words, especially when the word had been paired with three Subordinate Level exemplars during exposure (52.5% of the time). Thus, consistent with detecting word meanings cross-situationally, in this condition, in which both words came from the same semantic domain (e.g., dogs), children thought the one paired with 3 Subordinate Level exemplars (e.g. "puziv"), had a subordinate meaning (e.g., *Dalmatian*). The meaning selection pattern for the word that had been paired with 3 Basic Level exemplars was largely split between two types of meaning responses: many children thought this word had a Basic Level Meaning (33.3%), but many thought it had a Mutually Exclusive meaning (non-Dalmatian dogs) (41.0%). This suggests that some children adopted a contrastive meaning for the two words encountered in each exposure phase (e.g., one meaning *Dalmatian*, and the other meaning *non-Dalmatian dog*). But, surprisingly, children sometimes actually assigned a Subordinate Level meaning (25.6% of the time) to the 3 Basic Level Exemplar word (e.g., selecting only Dalmatians when tested on "ruften", a word that during the exposure phase had actually been paired with a variety of dogs). This would suggest that some children were confused about which word ("puziv" or "ruften") meant *Dalmatian*, or, perhaps children thought that both words had a Subordinate level meaning. At the end of this results section, we return to the important issue of how children paired meanings to the two words within each test block. But, regardless of those facts, it is clear that children in the Contrast at the Subordinate Level condition had a much stronger Subordinate Level meaning preference than children in the Contrast at the Basic Level condition, consistent with children finding contrasting meanings cross-situationally.

To test the reliability of these observations, we analyzed Subordinate Level Meaning responses as a function of Contrast Context (Contrast at the Basic Level vs. Contrast at the Subordinate Level) and Exemplar Type (3 Basic Level Exemplars vs. 3 Subordinate Level Exemplars). A mixed effect logistic regression is not appropriate for this data set because two of the cells from the Contrast at the Basic Level condition have essentially zero Subordinate Level Meaning responses (1 in each, resulting in the 0.025 proportions in Table 2A, row 1). Cells with rare events result in unstable logistic regressions due to poor maximum likelihood estimates (King & Zeng, 2001). We therefore computed an elogit transformation of subject and item means of Subordinate Meaning responses[3] and entered them into separate weighted linear models and report reliability by subject and item separately. The dependent variable was elogit of Subordinate Meaning Responses, with a random intercept given to each subject in the subject model and to each item in the item model. Two independent variables and their interaction served as fixed effects in a 2 × 2 design: Contrast Context (Contrast at the Basic Level vs. Contrast at the Subordinate Level) and Exemplar Type (3 Basic Level Exemplars vs. 3 Subordinate Level Exemplars). Deviation coding was used for both factors.

These analyses revealed a reliable interaction between Contrast Context and Exemplar Type by subjects ($\beta = -0.944$, $t(29) = -2.08$, $p = 0.046$) and by items ($\beta = -1.153$, $t(20) = -2.77$, $p = 0.012$). The interaction arose because there were more Subordinate Meaning Responses when a word had been paired with 3 Subordinate Level Exemplars as compared to 3 Basic Level exemplars, but only when words contrasted at the Subordinate Level (52.5% vs. 25.6% Subordinate responses see Table 2B, row 1) and not when words contrasted at the Basic Level (2.5% vs. 2.5% Subordinate responses see Table 2A, row 1). Indeed, simple effects revealed an effect of Exemplar Type for children whose words contrasted at the Subordinate Level, marginal by subjects ($\beta = -0.832$, $t(12) = -1.95$, $p = 0.074$) and significant by items ($\beta = -0.956$, $t(10) = -4.29$, $p = 0.016$) but no effect of Exemplar Type for children whose words contrasted at the Basic Level, by subjects ($\beta = 0.103$, $t(17) = 0.878$, $p = 0.392$), by items ($\beta = 0.196$, $t(10) = 0.606$, $p = 0.558$). This interaction and pattern of simple effects are expected if children are comparing novel word meanings cross-situationally and look for contrasting word meanings at the Subordinate level when the word meanings would have otherwise generated the same basic level meaning (e.g., both would have meant *dog*, as would have been the case in the Subordinate Contrast condition).

2.2.2. How children paired meanings within each exposure phase

Although the analyses we just reported are important for identifying children's general meaning preferences across conditions, they do not reveal the extent to which children differed in how they decided to assign meaning to the contrasting word pairs (e.g., "puziv" vs. "ruften") that they encountered during each block. For instance, when a particular child assigns a Subordinate Level meaning to one of the two words in a block, does he/she think the other word means *dog* (indicating a subset, inclusive, relation) or does he/she think the other word means *non-Dalmatian dog* (indicating two contrastive meanings). Since we did not pair our data in this way above, we have not yet answered this question, nor understood if children were internally consistent across the three blocks in assigning meaning. For instance, does a child consistently assign mutually exclusive meanings to each word pair on each block?

To answer these questions, we recoded our data at the word-pair (block) level. In particular for each block, we asked if the child thought the two words in that block: both had a Basic Level meaning (e.g., puziv = *dog*, ruften = *dog*), were Hierarchically related (e.g., puziv = *Dalmatian*; ruften = *dog*), were Contrastive (puziv = *Dalmatian*, ruften = *Non-Dalmatian dog*), or both had a Subordinate Level meaning (e.g., puziv = *Dalmatian*, ruften = *Dalmatian*). Strikingly, 26 of all 32 children were internally consistent, responding the same way across all usable blocks (N = 20) or the majority of useable blocks (N = 6). (See Appendix A for data from each subject at the trial level.) Given this consistency, we typed children based on their majority response and present the tallies for

[3] Specifically, for a given subject (or item) in a given condition, elogit = $\log((y + 0.5)/(N - y + 0.5))$, where y equals the number of subordinate meaning responses in that condition and N equals the total number of observations in that condition. See Barr (2008) for example. Subject corresponded to the individual child. Item corresponded to the individual nonsense word.

Table 3
Count of children based on preferred meaning response (Experiment 1).

	…for children whose words contrasted at the…	
Preferred meaning response…	Basic Level	Subordinate Level
Both Basic (e.g., puziv = *dog*, ruften = *dog*)	13	4
Contrastive (puziv = *Dalmatian*, ruften = *Non-Dalmatian dog*)	0	7
Both Subordinate (e.g., puziv = *Dalmatian*, ruften = *Dalmatian*)	0	2
Hierarchical (e.g., puziv = *Dalmatian*; ruften = *dog*)	0	0
No Type (Other)	3	3
Total	16	16

each type in Table 3 below, doing this separately for children in the Contrast at the Basic Level condition and the Contrast at the Subordinate Level condition.

Table 3 reveals the learning preferences of the children in this study. In particular, 13 of the 16 children who were given word pairs that contrasted at the Basic Level (such that, e.g., one was paired with 3 Dalmatians, the other with 3 different birds) picked Both Basic Level meanings at test (e.g., indicating that one word meant *dog,* and the other meant *bird*). Only three children did not do this and were instead untypable. (Children were untypable only because of too little data, due to having less than two usable blocks to assess their overall preference. A block was usable in this analysis if the child made no errors on the block, i.e., did not produce an error on either word in the block. See Appendix A.) Thus, for the Contrast at the Basic Level condition, children overwhelmingly thought both words had a basic level meaning.

In contrast, for those children given words that contrasted at the Subordinate Level (such that, e.g., one was paired with 3 Dalmatians, the other with 3 different dogs of different breeds), only 4 preferred Both Basic Level meanings (e.g., indicating that both words mean dogs generally). Seven others had Contrastive meanings for the pairs (e.g., indicating that one word meant *Dalmatian* and the other *non-Dalmatian dog*). Two children in this condition had Subordinate Level meanings for both words (e.g., indicating that both meant *Dalmatian*). Notably none of the children preferred Hierarchical meanings for the two words, such that one meant *Dalmatian* and the other meant *dogs* generally. This strongly suggests that children arrived at Subordinate meanings for words cross-situationally via a strategy of mutual exclusivity or contrast. And, it should be noted that the difference across the two Contrast Context conditions was statistically significant (i.e., 13 of 16 children had a basic level preference in the Contrast at the Basic Level condition whereas only 4 of the 16 did in the Contrast at the Subordinate Level condition, yielding a highly significant difference, $\chi^2(1) = 8.031$, $p = 0.0046$).

Thus, within the Subordinate Contrast condition, there was a split in children's ability, with 7 of the 16 children thinking the pairs of words within a block contrasted at the Subordinate level (7 with Contrastive Meaning, see Table 3). It is worth noting that this group did not distinguish itself from the rest of the children in terms of age (average age 51.0 vs. 49.1 months respectively, $t(14) = 0.468$, $p = 0.646$) nor gender (1 out of 7 were male vs. 5 out of 9 were male respectively, Fisher Exact test, $p = 0.145$). In short, neither age nor gender were a strong predictor of whether a child would arrive at contrastive meanings. While age and gender are unrelated to behavior, we speculate that children's memory and attention abilities may determine children's success in detecting lexical contrast cross-situationally, since as we mentioned above, children would need to notice that the to-be-learned words have the same meaning across situations and use this information to alter their meaning hypotheses.

2.2.3. Summary

The findings indicate that, as a group, 3-to-5-year-old children can arrive at subordinate level meanings for novel words from a series of simple ostensive labeling events, but only when the to-be-learned words have exemplars that come from the same semantic domain, resulting in words that have non-overlapping meanings contrasting at the Subordinate level. When the novel words came from different semantic domains, children overwhelmingly thought the words had basic level meanings. However, this cross-situational sensitivity to Contrast at the Subordinate Level was not present in all children, some children (4 of 16 in this condition) arrived at basic level meanings for all of the words. We believe that this pattern arises because simple ostensive labeling events lead children to assume the word meaning likely distinguishes itself at the basic level, and memory demands result in them not noticing across word occurrences that both words in a block have the same meaning.

We find little evidence consistent with children detecting suspicious coincidences cross-situationally. Specifically, in the Contrast at the Basic Level condition, subjects rarely (2.5% of the time) came up with Subordinate Level meanings, and instead thought the word had a Basic Level meaning (97.5% of the time). That is, they thought "puziv" meant *dogs* generally, even though every time children had heard "puziv" previously it had been paired with a Dalmatian. Subordinate interpretations would have been predicted by the suspicious coincidence account. Children did arrive at subordinate level meanings for such an exemplar sequence, but only when they were being asked to learn pairs of words that contrasted at the Subordinate level (Contrast at the Subordinate Level condition). Moreover, when children assigned a Subordinate Meaning to this word in this context, they always gave the second word (which had been paired with a random sample of dogs during exposure) a mutually exclusive meaning at test (e.g., *Non-Dalmatian dog*) suggesting children were comparing meanings of words, not tracking referential exemplar histories of words and computing likely semantic categories from those histories. If they had been doing the latter, they would have thought this other word meant *dogs* generally.

These findings contrast with the past findings from Xu and Tenenbaum (2007) who found that pairing a novel word with multiple subordinate level exemplars (e.g., 3 Dalmatians) leads children in this very same age range to have a near categorical preference for a subordinate (*Dalmatian*) meaning for the novel word. It is important to note though that the exposure conditions were very different in that past study, where three subordinate exemplars were presented simultaneously as a group to the child, labeled as a group of, e.g., "puzivs", and in the presence of a test array that contained a variety of other animals and objects. If their account did not depend on those circumstances, we should have seen a similar preference for subordinate meanings here when encountering 3 Subordinate Level Exemplars.

3. Experiment 2

Experiment 1 tested whether the traditional effects of learning words with subordinate meanings observed in the literature could still hold in situations in which children had to constrain their interpretations of novel word meanings cross-situationally, with individual exposures of each word to be learned with a single referent on the screen, rather than in the situation in which referents are simultaneously present and then labelled. Our findings indicated that inferential mechanisms such as contrast (Clark, 1987, 1988) and mutual exclusivity (Markman, 1987) that were previously identified to be at work in immediate local contexts are also at work during cross-situational comparison. In particular, we found that the mere act of contrasting novel word meanings at the subordinate level cross-situationally encourages children to arrive at subordinate level meanings. The influence of contrast was not as strong as what has been seen in past work (e.g., Gelman et al., 1989; Waxman et al., 1991), but this might be expected since we were asking children to contrast meanings across distinct situations.

As discussed in our introduction, the past literature has also shown that anchoring a novel word to a known basic level term allows children in this age range (3–5 years) to arrive at a subordinate level meaning for the novel word. For example, Waxman et al. (1991) Experiment 3, offered linguistic support for a subordinate level meaning (e.g., showing a Dalmatian and saying to the child "This is a dax. It's a kind of dog."; Exp. 3). Notably however, this test was done in a rich contrastive referential environment in which pairs of dogs were co-present, each labeled with a different novel word, and each highlighted with distinguishing characteristics ("Daxes help take care of sheep. Mipens find birds for hunters."). It is possible that such supportive information is needed to be present in the immediate referential context to appreciate linguistic anchoring. So, here we set out to isolate the contribution of linguistic anchoring itself in the absence of this additional support. Would encountering individual Dalmatians, each time labeled and anchored ("I see a puziv. It's a kind of dog") encourage subordinate level meanings, even in the absence of a co-present within-domain comparison exemplar (i.e., another dog) labeled with another term?

Our procedure was identical to Experiment 1 above, with the following changes. First, how Sally labeled exemplars during the exposure session was changed. Instead of hearing on each exposure trial "Oh wow, that's a puziv! I see a puziv. Do you see the puziv?", children would hear Sally say "Oh wow, that's a puziv! It's a kind of dog. I see a puziv! It's a kind of dog. Do you see the puziv?" Second, no within domain contrast was provided – the two terms in each exposure phase came from a different semantic domain (e.g., each time they heard "puziv" it was paired with a Dalmatian and children were told the puziv was a kind of dog, and each time they heard "ruften" it was paired with a Parrot and children were told the ruften was a kind of bird).

If linguistic anchoring on its own is sufficient to acquire subordinate terms, children should show a strong subordinate meaning preference for the novel words in this study at test (e.g., selecting only Dalmatians and not other dogs at test). This is a reasonable expectation because children in this age range surely know the basic level terms that we used as anchors (i.e., dog, bird, car, etc.) and they have in past studies shown an understanding of the "kind of" term. However, if further support is necessary to arrive at subordinate meanings, such that a confluence of supportive information is needed, we might expect a strong basic level preference for these words at test (e.g., selecting all dogs), despite being told that "puzivs" were a kind of dog.

3.1. Method

3.1.1. Participants

Twenty additional typically developing children between the age of 3 and 5 were recruited from preschools in the Philadelphia area (mean age = 48.6 months; Standard Deviation = 7.1 months; range = 36–59 months). All were native speakers of English. One additional child was tested but not included due to parental interference during the experiment. The children's parents provided informed consent in accordance with the IRB protocol of the University of Pennsylvania and were compensated with a small toy.

3.1.2. Stimuli

We used the same stimuli as in Experiment 1. Except that Exposure Phase utterances were re-recorded so that they mentioned the English basic level category, e.g., "Oh wow, that's a puziv! It's a kind of dog. I see a puziv! It's a kind of dog. Do you see a puziv?"

3.1.3. Design and procedure

The procedure was the same as Experiment 1. Children similarly received the practice phase and then three blocks of exposure and test. In each block, children were exposed to and tested on two novel words. Experiment 2 was modelled after the Contrast at the Basic Level condition of Experiment 1, in that children never learned two words belonging to the same domain (e.g., they never learned two words for dogs). However, unlike the Contrast at the Basic Level condition of Experiment 1, all words in Experiment 2 were assigned to the Subordinate Level condition, such that exemplars always belonged to the Subordinate Level category. Thus, in block 1, children were exposed to a word that referred to Dalmatians and another word that referred to parrots. In total, children

were exposed to six words belonging to the categories Dalmatian, parrot, basketball, police car, orange tabby cat and goldfish.

Critically, children were now given linguistic support for the subordinate-level meaning during the exposure phase. In particular, rather than hearing Sally say "Oh wow, that's an X! I see an X. Do you see the X?", children would hear Sally say "Oh wow, that's a X! It's a kind of (basic level category noun). I see a X! It's a kind of (basic level category noun). Do you see a X?" where X was replaced by a novel word in the singular form (e.g., "ruften"), and the basic level category noun was the basic level word for that category. For example, children would hear "Oh wow, that's a ruften! It's a kind of car. I see a ruften! It's a kind of car. Do you see a ruften?". The novel words we used referred to (in italics): Dalmatian – A kind of *dog*, Basketball – A kind of *ball*, Parrot – A kind of *bird*, Police car – A kind of *car,* Goldfish – A kind of *fish*, orange tabby cat – A kind of *cat*. After the sentence ended, there would be two seconds of silence before the next trial began. The test phase was as in Experiment 1.

Two experimental lists were created such that in List 2 the novel words in each block were swapped with the two categories in that block. Children were assigned to one of the two lists.

3.1.4. Predictions

The performance of children in Experiment 2 was to be compared to the performance of those children assigned to the Contrast at the Basic Level condition of Experiment 1. This is because in Experiment 2 all words belonged to different semantic domains, just like the Contrast at the Basic Level condition of Experiment 1. If children cannot use linguistic anchors to reason about levels of a semantic category, children in Experiment 2 should behave like the children in the Contrast at the Basic Level condition of Experiment 1, and thus should have an overwhelming bias to pick Basic Level Meanings at test. If, on the other hand, children can use linguistic anchors to reason about levels of a semantic category, children in this experiment should always pick the Subordinate Level meaning at test. For, if they had thought the nonsense word meant the basic level category, all of the exposure phase sentences would have been uninformative tautologies (e.g., "This is a dog. It is a kind of dog.")

3.2. Results and discussion

Following the coding procedure in Experiment 1, we excluded trials in which children selected incorrect referents. This procedure excluded 10.3% of all data. This means that children on average correctly learned the semantic domain of 89.7% of all words. We coded children's correct meaning responses in the same way as we did in Experiment 1, into one of three types: Subordinate Level Meaning (e.g., selecting only Dalmatians, and no other dogs), Basic Level Meaning (e.g., selecting all types of dogs, including Dalmatians), or Mutually Exclusive Meaning (e.g., selecting only non-Dalmatian dogs). Table 4 presents the proportion of each type of response.

As can be seen in Table 4, the dominant response pattern is the basic level, though subordinate meaning selections were at a relatively high rate (41.3% of all trials), similar to those children in Experiment 1 whose words contrasted at the Subordinate Level (52.5%, see 3 Subordinate Exemplars, Table 2B, row 1), and clearly different from children whose words contrasted at the Basic Level (2.5%, see 3 Subordinate Exemplars, Table 2A, row 1). Indeed as expected, children's elogit transformed Subordinate Level Meaning Responses in the present experiment were reliably different from children's elogit transformed Subordinate Level Meaning Responses in the Contrast at the Basic Level condition of Experiment 1 in this same condition, by subjects ($\beta = 1.566$, $t(33) = 2.735$, $p = 0.001$) and by items ($\beta = 2.155$, $t(5) = 11.21$, $p = 0.0001$).

The 41.3% Subordinate Meaning responses in the present experiment is striking given that the pairs of words in each block came

Table 4
Proportion of Type of Responses at Test. Dominant answer in bold.

	During Exposure
	3 Subordinate Level Exemplars
	"Puziv" "kind of dog"  "Puziv" "kind of dog"  "Puziv" "kind of dog" 
Proportion of responses at test for…	
Subordinate Level Meaning	0.413
Basic Level Meaning	**0.587**
Mutually Exclusive Meaning	0.000

from different semantic domains (one was paired with Dalmatians, the other with, e.g., Police Cars). One might expect that such a semantic contrast would result in a preference to think the two words had basic level meanings, e.g., *dog* and *car* respectively. Instead, children showed a preference for Subordinate level meanings similar to children in Experiment 1 who encountered word pairs that contrasted at the subordinate level (e.g., one pared with Dalmatians, the other with dogs generally). Thus, it is likely that linguistic anchoring (e.g., hearing that puzivs were a kind of dog, and ruftens were a kind of car, etc.) generated the relatively high rate of Subordinate Level meaning responses in this study.[4]

Importantly, just like in Experiment 1, children in Experiment 2 were highly consistent in their meaning responses at test. In particular, of the 20 children in this study, most responded the same way across all trials (N = 15) or on the majority of trials (N = 3). Only 2 were inconsistent. (See Appendix B for the data from all children at the trial level.) Given this consistency, we typed children based on their majority response. Unlike Experiment 1, this typing was done at the item rather than the block level because both items within a block did not differ in terms of condition in the present experiment. The tallies for each type appear in Table 5 below.

As can be seen in Table 5, a total of 8 of the 20 subjects behaved as if they understood what "kind of" meant and consistently preferred Subordinate Level meanings at test. In contrast though, 10 of the 16 subjects had a strong basic level bias: they indicated that e.g., "puziv" meant dogs generally, even though each exemplar during the exposure phase was a Dalmatian and described as "I see a puziv! It's a kind of dog"). Notably, the 'successful' children (i.e., the 8 who identified Subordinate Level meanings) did not differ from the rest of the children in terms of gender (4 out of 8 were male vs. 3 out of 12 were male respectively, $p = 0.173$ given Fisher's exact test) nor age (average age 51.6 and 47.5 months respectively, $t(18) = 1.643$, $p = 0.118$). While age and gender are unrelated to behavior, it is worth speculating why some children fail to arrive at subordinate level meanings. One possibility is that children who failed to generalize to the subordinate level were the ones who do not understand the meaning of the phrase "kind of". If some children correctly believe that "kind of" conveys a subset relation, and other children incorrectly believe it to convey an identity or similarity relation, this would lead to the results we obtained. Alternatively, it is possible that all children knew the meaning of "kind of" but some needed further contextual support to infer subordinate meanings (like those contextual factors present in the studies of Waxman and colleagues, e.g., Waxman et al., 1991, that found stronger effects of linguistic support). We leave this possibility for future research.

3.2.1. Summary

The findings indicate that some but not all children ages 3-to-5-years can arrive at subordinate level meanings cross-situationally when offered relatively minimal linguistic support ("It's a kind of dog."), even though the dominant response pattern in Experiment 2 is still at the basic level. On average, subordinate level meaning selections in this experiment were higher (41.3%) than subordinate level meaning selections in a similar condition from Experiment 1 (2.5%, in the Contrast at the Basic Level condition). This difference likely arose because the labeling events in Experiment 2 included linguistic support, in the form of anchoring to a known basic level term ("It's a kind of dog" and "It's a kind of car"). Further analysis showed that it was only a subset of children (8 of the 20) who consistently arrived at Subordinate Level meanings for these novel words. Notably, many (10 of 20) children instead selected basic level meanings (e.g., *dog, car*) at test for all or almost all words. This contrasts with past work (e.g., Waxman et al., 1991) in which nearly all children of similar ages (3 years of age) were able to arrive at subordinate level meanings in the presence of linguistic anchoring to a basic level term. As noted above however, the past work was done in very different referential circumstances in which pairs of dogs were co-present, each labeled with a different novel word, and each highlighted with distinguishing characteristics. It is likely then that children in this age range need a confluence of cues to consistently identify and distinguish subordinate level meanings, many of which were not present in the current experiment. In particular, simple ostensive labeling events in which the immediate visual environment contains only one visually present referent, like what was done here, likely encourages children to assume a word's meaning distinguishes itself at the basic level.

4. General Discussion

4.1. Summary and implications

In the work presented here, we explored some of the conditions under which children would adopt a subordinate level meaning for a word when referential exemplars are encountered individually across separate naming events. A large body of past work suggests that simple labeling environments such as these generate a basic level bias (thinking that words have basic level meanings like *dog*, or *bird*, rather than meanings such as *Dalmatian*, or *parrot*). This literature suggests that a non-basic level meaning is adopted only when a child is provided with a rich contextual environment that indicates the speaker is taking an alternative stance on the referent. Helpful contextual environments include situations in which groups of referents are co-present, the words referring to these groups are contrasted at the subordinate level, and distinguishing semantic information is highlighted for the child or when these

[4] It is also possible that the high rate of Subordinate Meaning responses in Experiment 2 was due to having all words paired with Subordinate level exemplars (Dalmatians, police cars, parrots, etc.). If this were the case though, one might expect this preference to emerge over the course of the experiment as children encounter more and more subordinate exemplars. However, this was not found to be the case: the proportion of subordinate meaning responses was 0.43 on the first block and 0.38 on the final block, which is the opposite of what would be expected if subordinate preferences emerged over time. Thus, it seems likely that linguistic anchoring to a basic level term (hearing "It's a kind of dog", "It's a kind of car", "It's a kind of bird", etc.) is driving the preference for subordinate meanings.

Table 5
Count of children based on preferred meaning response (Experiment 2).

Preferred meaning response	Total
Basic Level (e.g., puziv = dog, ruften = bird)	10
Subordinate Level (e.g., puziv = Dalmatian, ruften = Parrot)	8
No Type (Other)	2
Total	20

highlighting situations include linguistic anchoring to a known basic level term (e.g., Blewitt, 1983; Callanan, 1985, 1989; Clark & Wong, 2002; Gelman et al., 1989; Horton & Markman, 1980; Katz, Baker, & Macnamara, 1974; Markman, Horton, & McLanahan, 1980; Markman, 1987; Shipley et al., 1983; Waxman, 1990; Waxman et al., 1991, see also Waxman, Lynch, Casey, & Baer, 1997). Our findings are consistent with the expectations generated by this past literature: We do indeed observe a large basic level bias in our studies. However, we find evidence that the inferential mechanisms that were identified to be at work in more complex referential environments are also at work under our simpler labeling conditions cross-situationally. In particular, we found that the mere act of contrasting novel word meanings at the subordinate level cross-situationally encourages children to arrive at subordinate level meanings (see Experiment 1, Contrast at the Subordinate Level condition, as compared to Contrast at the Basic Level condition). Moreover, minimally introducing a linguistic anchor into the labeling event ("That's a puziv, it's a kind of dog") introduces near categorical arrival at subordinate meanings, at least in a minority of children in this age range (8 out of 20).

As noted in our introduction, more recent studies in the literature suggest that children can arrive at subordinate level meanings from a sequence of simple ostensive labeling events, even when additional sematic and linguistic support is not provided. In particular, Xu and Tenenbaum (2007) suggested that children could track the exemplars encountered from a series of individual labeling events, and compute the likelihood that these exemplars were, e.g., a random sample of a basic level category, e.g., detecting the suspicious coincidence that the dogs labeled as "puzivs" were Dalmatians, and not, as expected a variety of breeds. Yet, as we discussed in the introduction to this paper, this claim has not been adequately tested, since all positive findings to date with children have been under conditions in which all exemplars were presented at once and labeled as a group of "puzivs", which were presented in the context of co-present entities from the same semantic domain that were not labeled as "puzivs". Indeed, these experimental conditions contain many of the very special ingredients that the past literature suggests would be helpful for children in identifying a subordinate meaning. When in the present study we stripped away this potentially helpful information, and simply let children encounter individual Dalmatians labeled as "puzivs," we find no measurable ability for children to overcome their basic level preference (Experiment 1, Contrast at the Basic Level condition). At a minimum, our findings place boundary conditions on when children can use sampling statistics to infer word meanings, i.e., that such an inference can be drawn when exemplars are labeled as a group rather than individually, and when the semantic categories relevant for the inference are highlighted, rather than left unmentioned. Future work will need to determine if these factors, rather than sampling statistics, played the determinative role.

Even though future work is necessary, our present findings do suggest a plausible alternative explanation of the so-called suspicious coincidence effect. In particular, labeling a group of Subordinate exemplars with the same novel term ("puzivs") in the presence of other items not labeled as "puzivs," as was done in Xu and Tenenbaum (2007) invites visual inspection of the properties that these referents have in common and how they differ from the other visually co-present items, increasing the probability of noticing detailed common salient properties that are not defining of the basic level category (e.g. they all have spots) (see Caplan, under revision). This explanation does not involve reasoning about sampling statistics and instead pertains to attention and is similar to an explanation offered by Spencer et al. (2011), who found that adults do not arrive at a Subordinate Level interpretation when the group of exemplars are not presented simultaneously but instead presented serially, one second apart. Lewis and Frank (2016) replicated the Spencer et al. finding with adults but found that it does not completely eliminate the Subordinate Level preference as compared to seeing just one Subordinate Level exemplar labeled with a nonsense word. We note here, though, that an attentional account need not predict a complete elimination of a Subordinate preference for three exemplars when compared to a single exemplar condition. Indeed, modeling efforts (Caplan, under revision) appear to capture such effects using simple assumptions about attention rather than inferences about sampling statistics. We also note that our own efforts studying adults in similar cross-situational paradigms (Wang & Trueswell, in preparation) also point to an attentional account in adults, but crucially one that hinges on detecting semantic contrasts, like those examined in the present paper.

4.2. Learning in the moment and learning cross-situationally

The present findings dovetail nicely with other research on cross-situational word learning, which has studied learners' resolution of referential uncertainty (e.g., Medina, Snedeker, Trueswell, & Gleitman, 2011; Trueswell, Medina, Hafri, & Gleitman, 2013; Woodard et al., 2016; Yu & Smith, 2007; Yurovsky & Frank, 2015) rather than semantic uncertainty. This past work has led our group to speculate that what language learners are doing in the moment, as they hear speech, is forming a conjecture about the speaker's intended message, i.e., a semantic representation of the utterance itself, and it is only this information that is considered when associating a word form with possible semantic categories (Gleitman & Trueswell, 2018; Woodard et al., 2016). That is, although a child uses the referential context to form a conjecture about a speaker's intended message, the word forms are not associated with the referential context itself but rather with a semantic representation of the speaker's intended message.

Such a view greatly constrains how one would explain how identification of word meaning is accomplished in the present studies. We would like to propose that although the situational context might permit a wide range of ways of characterizing a particular referent (e.g., as a *Dalmatian, a dog, an animal, a pet,* etc.), the learner nevertheless makes an informed guess in the moment based on the current situational and conversational context. It is this conjecture that is entered into word learning. Word learners are doing this because they are trying to determine what a speaker has meant in the moment (i.e., the speaker's message). When a word is heard again, that meaning is retrieved from memory and evaluated with respect to the new context. A speaker who is using single words to distinguish different animals (and not, e.g., the same word plus a linguistic device to modify it – "a jub" vs. "a spotted jub") leads to hypothesized meanings at a particular level and not others, in the way so described above. The word form is associated with this hypothesized meaning and not, for instance, a range of hypothesized meanings, nor some 'unanalyzed' perceptual representation of the referent (e.g., what that particular dog looked like in that context). This latter view would likely have to be proposed for any account that lets learners reason about referential exemplars across word learning instances, such as the Xu and Tenenbaum (2007) account. This is because the learner would need to remember all the possible semantic categories each exemplar belonged to if one is to determine which of these categories deviate from what would be expected under conditions of random sampling. It is important to note that the Xu and Tenenbaum (2007) account of how learners deal with semantic ambiguity is computationally analogous to unrestricted theories of cross-situational word learning, which assume that learners will hold in abeyance any commitment to a word meaning while evidence accumulates across multiple occurrences of a word. Indeed, the Bayesian account of referential uncertainty (Frank, Goodman, & Tenenbaum, 2009) is a computational variant of the Bayesian account of semantic uncertainty in Xu and Tenenbaum (2007).

We suggest that like the findings on referential ambiguity, the present findings on semantic ambiguity are inconsistent with such a view of how learners deal with semantic uncertainty. Rather than patiently accumulating evidence, learners select a single meaning in the moment, and this selection alone enters into word learning. We agree with others that complex inferential machinery must be at work to identify meanings in the moment, e.g., sensitivity to pragmatic aspects of language use (such as what is being contrasted across local utterances, as was the case in Experiment 1 above). At issue here instead is what memory representations learners have access to and could plausibly consider when a word is uttered. Much of these issues will need to be addressed in future work that examines the time course of learning and, in particular, examines the memory representations accessed when a word is encountered again under different learning conditions. But, it is particularly telling that our present findings demonstrate that children successfully compared word meanings cross-situationally. In particular, in Experiment 1, when simple labeling events pointed to two words having the same Basic Level meaning (e.g., both words support the meaning *dog,* in the Subordinate Contrast condition), there was a significant increase in likelihood that children would try to identify two meanings contrasting at the subordinate level (deciding that e.g., one novel word meant Dalmatian, and the other meant non-Dalmatian dogs). This suggests children had memory access to conjectured word meanings. The inability to detect suspicious coincidences, on the other hand, suggests they did not have memory access to unconjectured meanings nor to 'unanalyzed' perceptual representations of past referents.

4.3. Why the use of familiar referents?

In this context, it is notable that like Xu and Tenenbaum (2007), we chose to use visual referents with which children are likely already familiar. That is, cars, birds, dogs, etc. were used as stimuli, rather than unfamiliar/novel objects and animals. This was done because like the work of Xu and Tenenbaum (2007), the central interest of the present work was to ask how children map novel words onto existing semantic knowledge that they have of the world, including hierarchical semantic knowledge. We were not exploring how children form new concepts nor how they might arrange these concepts hierarchically based on perceptual or linguistic experience. Current theories of word learning generally agree that word learning can occur under such conditions, making word learning a so-called "mapping problem" (e.g., Gleitman & Trueswell, 2018). Much more debated is whether all word learning occurs in this fashion, such that children must have a concept and conceptual organization in place prior to identifying a word's meaning, or if the linguistic act of labeling itself can also, under some conditions, invite the formation of a new concept. There is evidence offered in support of both views. In particular, there is evidence that successful word learning in infants (12-month-olds) may require the preformation of categories (e.g., Jenkins et al., 2015; Matlen, Fisher, & Godwin, 2015; Pomiechowska & Gliga, 2019; Sloutsky & Fisher, 2004). There is however other evidence suggesting that hearing novel words may trigger category formation itself in infants (e.g., see Ferguson & Waxman, 2017, and references therein). The present work does not speak to this debate, but the debate influences the scope of the conclusions we might draw from our findings. If one assumes (as we do here) that across development language only or primarily maps onto preexisting conceptual knowledge (e.g., Clark, 2017; Mandler, 2004), then our present observations and conclusions might apply to earlier stages of word learning. Such a view would suggest that children apply the same word learning strategies (contrast, etc.) early on, but necessarily operate on more limited knowledge. For example, it goes without saying that a child who has never seen dogs before is unlikely to behave the same way in our study as a child who has experience with dogs. However, if instead we first expose children to novel entities in a nonlinguistic context, and offer them an opportunity to develop perceptual and/or functional knowledge that mentally organizes these entities, then the learning of novel words that label these newly acquired concepts might yield results similar to those reported here, even in very young children who are acquiring their first words (e.g., 12 months). We leave these issues for future work.

4.4. Where does the basic level bias come from?

From the view just espoused, a more unified account emerges for why infants and young children arrive at particular levels of meaning for words. A language novice, who is not already equipped with a large vocabulary and a syntactic grammar, is forced to learn their first words from observation of their immediate perceptual surroundings, which radically limits their initial vocabulary to words whose meanings are highly imageable (Gillette, Gleitman, Gleitman, & Lederer, 1999; Gleitman, 1990). Words with abstract or non-imageable meanings, be they nouns or verbs, are not easily learned by observation alone, even by older children and adults who possess these conceptual categories (for adults, see Gillette et al., 1999; for international adoptees learning English, see Snedeker, Geren, & Shafto, 2007). With respect to the current discussion, many (but not all) semantic categories dubbed "superordinate" within semantic hierarchies are abstract and/or non-imageable. Starting with Rosch and Mervis (1975), many have noted that superordinate categories are characterized by having low within-category physical resemblance (e.g., *animal*, *object*) and fewer defining attributes, which often means their attributes are functional rather than perceptual (e.g., *furniture, vehicle, weapon*). We believe that this fact, and perhaps this fact alone, explains why superordinates do not heavily populate the early vocabulary of language novices, be they children learning their first language or older children/adults learning a second language. That is, observation of the referent world makes it hard to predict when someone meant to say "furniture", "vehicle" or many other superordinate terms (see Gillette et al. for a general discussion on the difficulty of learning abstract words from observation alone). It follows from such a view that if children's first words are learned from observation of the referent world, many early word meanings should be perceptually defined, in the sense that they distinguish objects from each other. This has been suggested in the literature and has found experimental support, such that preferred word meanings usually end up being concrete object categories whose definition relies on distinguishing perceptual object characteristics such as shape (e.g., Landau, Smith, & Jones, 1988).

Under this view, observational word learning "pushes" the first acquired vocabulary down to the so-called "basic level". But, what about the paucity of *subordinates* in a learner's early vocabulary? One cannot appeal to observational factors alone here, for most subordinate terms are just as imageable or perhaps even more imageable than so-called basic level terms (consider again "dog" vs. "Dalmatian"; "bird" vs. "parrot"). Here we think our finding that local semantic contrast matters for identifying word meanings might play some role. In particular, if one assumes that speakers tend to select terms that contrast at the same semantic level for use within a local discourse context (e.g., we are talking about types of dogs; or types of animals), then a child who already knows the meaning of some basic level terms has an opportunity to identify the meanings of other basic level terms that they don't already know. For example, it seems more likely that someone will say "I see a dog and a cat" not "a dog and a Manx" nor "a poodle and a cat." Thus, given typical term choice patterns of speakers, children who know some basic level terms will likely learn more basic level terms. Moreover, these opportunities for basic-level word learning would be further exaggerated if speakers preferentially use basic level terms with children (e.g., Shipley et al., 1983). In general, however, these issues of conversational patterns and resulting word learning opportunities are left for future work. It does follow though from such a view that the learning of non-basic level terms would likely require additional situational and/or linguistic support not present in typical ostensive labeling contexts. And indeed, an earlier literature which we reviewed above, explored and found confirmatory evidence for this conclusion (e.g., Blewitt, 1983; Callanan, 1985, 1989; Clark & Wong, 2002; Gelman et al., 1989; Horton & Markman, 1980; Katz et al., 1974; Markman et al., 1980; Markman, 1987; Shipley et al., 1983; Waxman, 1990; Waxman et al., 1991). In the present studies we find that the same is true for cross-situational comparison: To the extent that children identify subordinate level meanings from a series of simple ostensive labeling events, it is done when children notice that their interlocutor couldn't possibly be using two different words that have the same meaning (Experiment 1) or when this interlocutor provides explicit linguistic support that they are mature enough to interpret and understand (Experiment 2). In this way, further information that they have learned about the language is brought to bear on the immediate situation to identify a plausible meaning. This meaning is stored in memory with the word form, recalled at its next occurrence, and used yet again to hypothesize the speaker's intended meaning.

Author note

This work was funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development Grant 1-R01-HD-37507 to JCT. The original data from the experiments as well as an analysis script reported in this paper can be found at https://osf.io/hnkq7/. We thank Lila Gleitman, Alex de Carvalho, Anna Papafragou, and Victor Gomes for their helpful comments. We also thank the area preschools, parents, and children who participated in these studies.

Appendix A. Meaning responses at test from each subject in Experiment 1

For each Block, children were exposed to two novel words, one paired with 3 Basic Level Exemplars (e.g., 3 Different Dogs) and the other with 3 Subordinate Level Exemplars (3 Dalmatians). Coded meaning responses at test appear for each word in each block. The first code is for the word paired with 3 Basic Level Exemplars, the second for the word with 3 Subordinate Level Exemplars. Codes: BL = Basic Level meaning; SL = Subordinate Level meaning; ME = Mutually Exclusive meaning; E = Error. Overall Type indicates the dominant response pattern for that subject (see Experiment 1 Results).

Condition	Subject	Block 1	Block 2	Block 3	Overall Type
Contrast at Basic Level	1	BL, BL	BL, BL	BL, BL	Both Basic
	2	BL, BL	BL, BL	BL, BL	Both Basic
	3	BL, BL	BL, BL	BL, BL	Both Basic
	4	BL, BL	BL, BL	BL, BL	Both Basic
	5	BL, BL	BL, BL	BL, BL	Both Basic
	6	BL, BL	BL, BL	BL, E	Both Basic
	7	BL, BL	BL, BL	BL, BL	Both Basic
	8	BL, BL	BL, BL	BL, BL	Both Basic
	9	BL, BL	BL, BL	BL, BL	Both Basic
	10	SL, BL	BL, BL	BL, BL	Both Basic
	11	BL, BL	BL, BL	BL, E	Both Basic
	12	BL, BL	BL, BL	E, BL	Both Basic
	13	BL, BL	BL, BL	E, E	Both Basic
	14	BL, BL	BL, E	E, E	Untypeable
	15	BL, BL	E, BL	BL, E	Untypeable
	16	E, E	E, E	E, SL	Untypeable
Contrast at Subordinate Level	17	BL, BL	BL, BL	BL, BL	Both Basic
	18	BL, BL	BL, BL	BL, BL	Both Basic
	19	BL, BL	BL, BL	E, E	Both Basic
	20	BL, BL	SL, BL	BL, BL	Both Basic
	21	SL, SL	SL, SL	ME, ME	Both Subordinate
	22	SL, SL	SL, SL	SL, SL	Both Subordinate
	23	ME, SL	ME, BL	ME, SL	Contrastive
	24	ME, SL	ME, SL	E, E	Contrastive
	25	ME, SL	ME, SL	ME, SL	Contrastive
	26	ME, SL	ME, SL	ME, SL	Contrastive
	27	ME, SL	SL, E	ME, SL	Contrastive
	28	ME, SL	SL, ME	BL, SL	Contrastive
	29	SL, ME	ME, SL	SL, SL	Contrastive
	30	BL, BL	E, BL	E, E	Untypeable
	31	BL, BL	E, BL	E, E	Untypeable
	32	E, E	E, E	E, E	Untypeable

Appendix B. Meaning responses at test from each subject in Experiment 2

For each Block, children were exposed to two novel words, both paired with 3 Subordinate Level exemplars from different semantic domains (e.g., one paired with 3 Dalmatians, the other with 3 Parrots). Coded meaning responses at test appear for each word in each block, corresponding to each word. Codes: BL = Basic Level meaning; SL = Subordinate Level meaning; ME = Mutually Exclusive meaning; E = Error. Overall Type indicates the dominant response pattern for that subject (see Experiment 2 Results).

Subject	Block 1	Block 2	Block 3	Overall Type
1	BL, BL	BL, BL	BL, BL	Basic Level Meaning
2	BL, BL	BL, BL	BL, BL	Basic Level Meaning
3	BL, BL	BL, BL	BL, BL	Basic Level Meaning
4	BL, BL	BL, BL	BL, BL	Basic Level Meaning
5	BL, BL	BL, BL	E, E	Basic Level Meaning
6	BL, BL	BL, E	BL, BL	Basic Level Meaning
7	BL, BL	BL, E	BL, BL	Basic Level Meaning
8	BL, BL	BL, E	E, BL	Basic Level Meaning
9	BL, E	BL, BL	BL, BL	Basic Level Meaning
10	E, E	BL, BL	BL, BL	Basic Level Meaning
11	SL, SL	SL, SL	SL, SL	Subordinate Level Meaning
12	SL, SL	SL, SL	SL, SL	Subordinate Level Meaning
13	SL, SL	SL, SL	SL, SL	Subordinate Level Meaning
14	SL, SL	SL, SL	SL, SL	Subordinate Level Meaning
15	SL, SL	SL, SL	SL, E	Subordinate Level Meaning
16	SL, SL	SL, SL	BL, BL	Subordinate Level Meaning
17	SL, SL	BL, SL	SL, BL	Subordinate Level Meaning
18	SL, BL	SL, SL	SL, SL	Subordinate Level Meaning
19	BL, BL	BL, E	E, E	Untypable
20	BL, SL	BL, E	E, E	Untypable

References

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language, 59*(4), 457–474.
Blewitt, P. (1983). Dog versus Collie: Vocabulary in speech to young children. *Developmental Psychology, 19*(4), 602–609.
Callanan, M. A. (1985). How parents label objects for young children: The role of input in the acquisition of category hierarchies. *Child Development, 56*(2), 508–523.
Callanan, M. A. (1989). Development of object categories and inclusion relations: Preschoolers' hypotheses about word meanings. *Developmental Psychology, 25*(2), 207.
Caplan, S. (under revision). Word learning as category formation. Unpublished manuscript.
Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.). *Mechanisms of language acquisition* (pp. 1–33). Hillsdale, NJ: Lawrence Earlbaum Assoc.
Clark, E. V. (1988). On the logic of contrast. *Journal of Child Language, 15*(2), 317–335.
Clark, E. V. (2017). Semantic categories in acquisition. *Handbook of categorization in cognitive science* (pp. 397–421). Elsevier.
Clark, E. V., & Wong, A. D. W. (2002). Pragmatic directions about language use: Offers of words and relations. *Language in Society, 31*(2), 181–212.
Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development, 11*(3), 356–373.
Diesendruck, G., & Shatz, M. (2001). Two-year-olds' recognition of hierarchies: Evidence from their interpretation of the semantic relation between object labels. *Cognitive Development, 16*(1), 577–594.
Ferguson, B., & Waxman, S. (2017). Linking language and categorization in infancy. *Journal of Child Language, 44*(3), 527–552.
Frank, M. C., Goodman, N. D., & Tenenbaum, J. B. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science, 20*, 578–585.
Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology, 40*(2), 177.
Gelman, S. A., Wilcox, S. A., & Clark, E. V. (1989). Conceptual and lexical hierarchies in young children. *Cognitive Development, 4*(4), 309–326.
Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition, 73*(2), 135–176.
Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1*(1), 3–55.
Gleitman, L. R., & Trueswell, J. C. (2018). Easy words: Reference resolution in a malevolent referent world. *Topics in Cognitive Science*.
Golinkoff, R. M., Shuff-Bailey, M., Olguin, R., & Ruan, W. (1995). Young children extend novel words at the basic level: Evidence for the principle of categorical scope. *Developmental Psychology, 31*(3), 494–507.
Gottfried, G. M., & Tonks, S. J. (1996). Specifying the relation between novel and known: Input affects the acquisition of novel color terms. *Child Development, 67*(3), 850–866.
Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children.* Baltimore, MD: Paul H. Brookes Publishing.
Horton, M. S., & Markman, E. M. (1980). Developmental differences in the acquisition of basic and superordinate categories. *Child Development, 8*(3), 708–719.
Jenkins, G. W., Samuelson, L. K., Smith, J. R., & Spencer, J. P. (2015). Non-Bayesian noun generalization in 3-to 5-year-old children: Probing the role of prior knowledge in the suspicious coincidence effect. *Cognitive Science, 39*(2), 268–306.
Katz, N., Baker, E., & Macnamara, J. (1974). What's in a name? A study of how children learn common and proper names. *Child Development, 45*(2), 469–473.
King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis, 9*(2), 137–163.
Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*(3), 299–321.
Lewis, M. L., & Frank, M. C. (2016). Understanding the effect of social context on learning: A replication of Xu and Tenenbaum (2007b). *Journal of Experimental Psychology: General, 145*(9), e72.
Markman, E. M. (1987). How children constrain the possible meanings of words. In U. Neisser (Ed.). *Concepts and conceptual development: Ecological and intellectual factors in categorization.* Cambridge: Cambridge University Press.
Markman, E. M., Horton, M. S., & McLanahan, A. G. (1980). Classes and collections: Principles of organization in the learning of hierarchical relations. *Cognition, 8*(3), 227–241.
Mandler, J. M. (2004). Thought before language. *Trends in Cognitive Sciences, 8*(11), 508–513.
Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences, 108*(22), 9014–9019.
Matlen, B. J., Fisher, A. V., & Godwin, K. E. (2019). The influence of label co-occurrence and semantic similarity on children's inductive generalization. *Front. Psychol. 6*. https://doi.org/10.3389/fpsyg.2015.01146.
Pomiechowska, B., & Gliga, T. (2019). Lexical acquisition through category matching: 12-month-old infants associate words to visual categories. *Psychological Science, 30*(2), 288–299.
Quine, W. V. O. (1960). *Word and object (studies in communication).* New York and London: Technology Press of MIT.
Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7*(4), 573–605.
Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*(3), 382–439.
Shipley, E. F., Kuhn, I. F., & Madden, E. C. (1983). Mothers' use of superordinate category terms. *Journal of Child Language, 10*(3), 571–588.
Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition, 61*(1–2), 39–91.
Sloutsky, V. M., & Fisher, A. V. (2004). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General, 133*(2), 166.
Snedeker, J., Geren, J., & Shafto, C. L. (2007). Starting over: International adoption as a natural experiment in language development. *Psychological Science, 18*(1), 79–87.
Spencer, J. P., Perone, S., Smith, L. B., & Samuelson, L. K. (2011). Learning words in space and time probing the mechanisms behind the suspicious-coincidence effect. *Psychological Science, 22*(8), 1049–1057.
Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology, 66*(1), 126–156.
Wang, F. H., & Trueswell, J. C. Being suspicious of suspicious coincidences: The case of learning subordinate word meanings (in preparation).
Waxman, S. R. (1990). Linguistic biases and the establishment of conceptual hierarchies: Evidence from preschool children. *Cognitive Development, 5*(2), 123–150.
Waxman, S. R., & Hatch, T. (1992). Beyond the basics: Preschool children label objects flexibly at multiple hierarchical levels. *Journal of Child Language, 19*(1), 153–166.
Waxman, S. R., Lynch, E. B., Casey, K. L., & Baer, L. (1997). Setters and samoyeds: The emergence of subordinate level categories as a basis for inductive inference. *Developmental Psychology, 33*(6), 1074–1090.
Waxman, S. R., Shipley, E. F., & Shepperson, B. (1991). Establishing new subcategories: The role of category labels and existing knowledge. *Child Development, 62*(1), 127–138.
Woodard, K., Gleitman, L. R., & Trueswell, J. C. (2016). Two-and three-year-olds track a single meaning during word learning: Evidence for Propose-but-verify. *Language Learning and Development, 12*(3), 252–261.
Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review, 114*(2), 245–272.
Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science, 18*(5), 414–420.
Yurovsky, D., & Frank, M. C. (2015). An integrative account of constraints on cross-situational learning. *Cognition, 145*, 53–62.