COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science (2017) 1–15

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12516

Learning Object Names at Different Hierarchical Levels Using Cross-Situational Statistics

Chi-hsin Chen,[a] Yayun Zhang,[b] Chen Yu[b]

[a]*Department of Otolaryngology-Head and Neck Surgery, The Ohio State University*
[b]*Department of Psychological and Brain Sciences, Indiana University*

Received 17 December 2016; received in revised form 25 May 2017; accepted 31 May 2017

Abstract

Objects in the world usually have names at different hierarchical levels (e.g., *beagle*, *dog*, *animal*). This research investigates adults' ability to use cross-situational statistics to simultaneously learn object labels at individual and category levels. The results revealed that adults were able to use co-occurrence information to learn hierarchical labels in contexts where the labels for individual objects and labels for categories were presented in completely separated blocks, in interleaved blocks, or mixed in the same trial. Temporal presentation schedules significantly affected the learning of individual object labels, but not the learning of category labels. Learners' subsequent generalization of category labels indicated sensitivity to the structure of statistical input.

Keywords: Statistical word learning; Word generalization; Category learning; Hierarchical labels; Temporal presentation schedules

1. Introduction

There are multiple levels of ambiguity in most word learning scenarios. As Quine's (1960) referential uncertainty problem illustrates, the world offers a seemingly infinite number of possible word-meaning mappings. Every time a learner hears a novel word, it is often ambiguous whether the word refers to one of the objects present in view, to part of an object, to its property or movement, or to the scene as a whole. Even though the referential uncertainty problem was originally raised as a philosophical puzzle, it has a huge influence on the conceptualization of word learning problems in language acquisition. From a cognitive perspective, we can further decompose the referential uncertainty

Correspondence should be sent to Chi-hsin Chen, Department of Otolaryngology-Head and Neck Surgery, The Ohio State University, 915 Olentangy River Rd, Columbus, OH 43212. E-mail: Chi-Hsin.Chen@osumc.edu

problem into two subsequent learning tasks. As shown in Fig. 1, when a language learner hears "here is a *modi*," there are many objects in the scene that serve as potential referents of the novel word *modi*. Thus, the learner needs to first discover which object goes with the word. This *word-referent mapping* problem has long been the focus of empirical research (e.g., Baldwin, 1993; Markman & Wachtel, 1988; Smith, Jones, & Landau, 1996; Tomasello & Akhtar, 1995). After the learner successfully links the word *modi* with the object "dog" in the scene, she needs to infer the meaning of *modi* from the observed object(s). More specifically, does the word *modi* refer to the particular dog in the scene, one breed of dogs (e.g., beagles), all dogs, or all animals? This example also demonstrates that an object can have names at different hierarchical levels (e.g., *beagle*, *dog*, *animal*) and a single label may be mapped to more than one object (e.g., the word *dog* can be used to refer to different instances of dogs).

Numerous studies have shown that young children, and even adults to some degree, have a tendency to assume that one object has only one name (e.g., Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992; Markman & Wachtel, 1988). However, it has also been documented that children as young as 2 years of age can go beyond one-to-one mappings and comprehend and produce labels at different hierarchical levels (e.g., *cat* and *animal*, Clark & Svaib, 1997). To do so, young language learners need to solve both word-referent mapping and word-meaning mapping problems.

Cross-situational learning has been proposed as a solution to early word learning. The idea is that language learners use word-object co-occurrences across different situations to learn the mappings between words and their potential referents/meanings. A growing body of research suggests that humans are able to use co-occurrence information gathered across situations to learn nouns, verbs, and adjectives (e.g., Akhtar & Montague, 1999; Childers & Paik, 2009; Scott & Fisher, 2012; Smith & Yu, 2008; Suanda & Namy, 2012; Vlach & Johnson, 2013; Yu & Smith, 2007) and to simultaneously acquire nouns and verbs (Monaghan, Mattock, Davies, & Smith, 2015). A few recent studies have shown

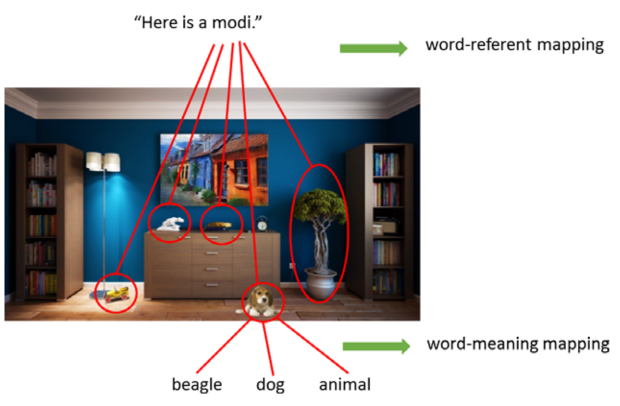

Fig. 1. Word-referent mapping and word-meaning mapping problems. In order to successfully learn a novel label, learners not only need to figure out the target referent of a label from many potential candidates, but also need to infer whether or how the label can be used to refer to other objects.

that adults are capable of using cross-situational statistics to map two labels to a same object (e.g., Benitez, Yurovsky, & Smith, 2016; Poepsel & Weiss, 2014), map two objects to a same label (Yurovsky, Yu, & Smith, 2013), and even map two labels to two objects (Kachergis, Yu, & Shiffrin, 2012). These studies indicate the possibility of using cross-situational statistics to learn words at different hierarchical levels, as learning hierarchical labels involves the mappings between one object and multiple words (e.g., an object mapped to the words *beagle* and *dog*) as well as the mappings between one word and multiple objects (e.g., the word *dog* mapped to different dog instances). Recently, Gangwani, Kachergis, and Yu (2010) conducted a cross-situational learning study, in which each learning trial contained two novel objects and three novel words (Experiment 1). Two of the words in a trial were mapped to only one object each and the third word was mapped to both objects. Participants were able to learn both one-word-to-one-object and one-word-to-many-objects mappings. This study provides initial evidence suggesting that cross-situational statistics can be used to acquire hierarchical labels.

Learning the meanings of object labels is more than just attaching a novel word to a particular referent (or referents) encountered when the word is heard. It involves an understanding of how broadly the word can be generalized beyond the encountered referent(s). In a word generalization study, Xu and Tenenbaum (2007) found that adults and children are sensitive to the range spanned by the objects they saw (e.g., three different dogs or three different animals) when hearing a label. Their subsequent generalization of that label is consistent with the span of the encountered instances (e.g., to other dogs or to other animals).

To date, however, the word learning literature has mainly focused on either the word-referent mapping problem (e.g., cross-situational learning studies) *or* the word-meaning mapping problem (e.g., word generalization studies), but not both (see Chen, Gershkoff-Stowe, Wu, Cheung, & Yu, 2016, for a recent exception). To have a more complete picture of the word learning processes, the present paper investigates adults' cross-situational learning of labels belonging to different hierarchies and their subsequent generalization of category labels. Specifically, we explore adults' ability to use co-occurrence information to learn both individual labels and category labels for objects belonging to different categories. Furthermore, we test whether learners generalize category labels to novel instances and, if so, whether their generalization scope is consistent with the span of the objects encountered during learning phase. These tasks require learners to simultaneously solve both word-referent and word-meaning mapping problems through a brief training session.

The second goal of the current project is to examine the effect of temporal arrangement on the learning of individual and category labels. One challenge of learning multiple labels for the same object is that different labels may compete with each other and interfere with the learning of the other. Benitez et al. (2016) found that when learning two labels for a same object, inter-item competition is reduced when the two labels are presented in completely separated blocks. Similar benefits of blocked designs have been found in several other cross-situational word learning studies (e.g., Vlach & Johnson, 2013; Yurovsky et al., 2013). On the other hand, some previous category learning studies have shown that interleaved object presentation promotes category learning and generalization (e.g., Kang & Pashler, 2012; Kornell & Bjork, 2008; Wahlheim, Dunlosky, & Jacoby, 2011; though see also Carvalho & Goldstone, 2015; Vlach & Kalish, 2014). These two lines of studies suggest that word learning and category learning may benefit from different types of temporal presentations. Therefore, one question to ask is whether temporal arrangement of label presentations differentially affects cross-situational learning of labels for individual objects and labels for categories. To test this possibility, we studied individual and category label learning using three different temporal arrangements: completely separated blocks, interleaved blocks, and mixed in the same trial. From a word learning perspective, having individual and category labels presented in blocks may minimize inter-item competition. However, from a category learning perspective, blocked designs may not be optimal for learning category labels. Presenting individual and category labels in different blocked schedules (i.e., separated vs. interleaved) allows us to test whether these two types of labels have to be completely non-overlapping or whether they can occur in close proximity. On the other hand, in natural language learning environments, labels heard in the same context do not always belong to the same hierarchical level. For example, in the sentence "I left my mug on the table," the word *mug* is a subordinate-level word, which belongs to the basic-level category of *cup*, while the word *table* is a basic-level word. Presenting individual and category labels in the same trial allows us to simulate this type of learning scenario.

In the following experiment, participants went through cross-situational learning trials to learn labels for different categories and labels for individual objects. These two types of labels were presented in three temporal arrangements: completely separated blocks (Separated condition), interleaved blocks (Interleaved condition), or mixed in the same trial (Mixed condition). In addition to examining participants' ability to learn hierarchical labels presented in different temporal arrangements, we also investigated whether and how learners generalize newly acquired category labels to novel instances.

2. Method

2.1. Participants

Participants were 79 undergraduate students (56 females, mean age: 18.93, SD = 1.33) at Indiana University who received course credits for volunteering. They were randomly assigned into one of three conditions: Separated condition (n = 27), Interleaved condition (n = 26), or Mixed condition (n = 26).

2.2. Stimuli and design

2.2.1. Stimuli

Unlike most cross-situational studies that used novel objects as visual stimuli, previous studies on how input structure affected word generalization mainly used real objects organized around different taxonomic categories (e.g., Spencer, Perone, Smith, & Samuelson, 2011; Xu & Tenenbaum, 2007). Following this line of work, the training stimuli used in the current study were real objects organized around different basic-level categories. This design allowed us to pre-determine category boundaries and examine whether participants over-generalized category labels to objects at different hierarchical levels.

Two sets of visual training stimuli were used. Each set consisted of 16 pictures of real objects that belonged to four different basic-level categories, two natural kind (e.g., dogs and peppers) and two artifact categories (e.g., chairs and cups, see Fig. 2A). The items were selected based on two criteria: (a) being a clear member of a training category, and (b) easily distinguishable from other members of the same category. Two naïve research assistants were presented with the pictures, one at a time, and asked to come up with a basic-level label for each item. The agreement in responses was 96.9%. Items in a category were also paired with each other. The research assistants were asked to judge how easy it was to perceptually distinguish between the items in each pair on a scale of 1 (not distinguishable) to 7 (easily distinguishable). The mean rating across all pairs was 6.92 (range: 6-7). The training auditory stimuli included 20 pseudo-words that followed the phonological rules of English. Of the 20 words, 16 were used as individual labels, each mapped to only one object (e.g., *vamy* in Fig. 2A), and 4 were category labels, each mapped to four objects (e.g., *zorch* in Fig. 2A).

For each training set, 16 additional objects were used at test as generalization items (four objects per category). Half of the objects were novel instances from the training basic-level categories (e.g., a German shepherd and a dachshund for the dog category). The other half of the objects were novel instances that belonged to the same superordinate-level category, but came from a different basic-level category as the training objects. For example, one of the items was a rabbit and another was a pig. Both objects belong to the superordinate *animal* category, but they are not dogs.

2.2.2. *Training session*

There were two types of training trials, individual label learning trials and category label learning trials. In each individual label learning trial, participants saw three objects, each from a different category, and heard their individual labels presented in random order (e.g., Trials 1, 9 and 13 in Fig. 2B). Within each trial, word-object mappings were ambiguous. However, across trials, each individual label co-occurred consistently with only one object. For example, given the information in Trial 1, learners could not tell which object was mapped to the word *vamy*. However, in Trial 13, if participants remembered having heard the word *vamy* while seeing the beagle, they should be able to infer that *vamy* was mapped to the beagle. Similarly, in each category label learning trial, participants were not given the information of which object was mapped to which word. However, every time they heard the word *zorch* (Trials 33, 37, and 48 in Fig. 2B), they saw a dog in the same trial. Keeping track of word-object co-occurrences should allow participants to learn the mappings between objects and category labels as well.

Over the training session, each object occurred nine times, six times with their individual label and three times with the category label. This design yielded a total of 48

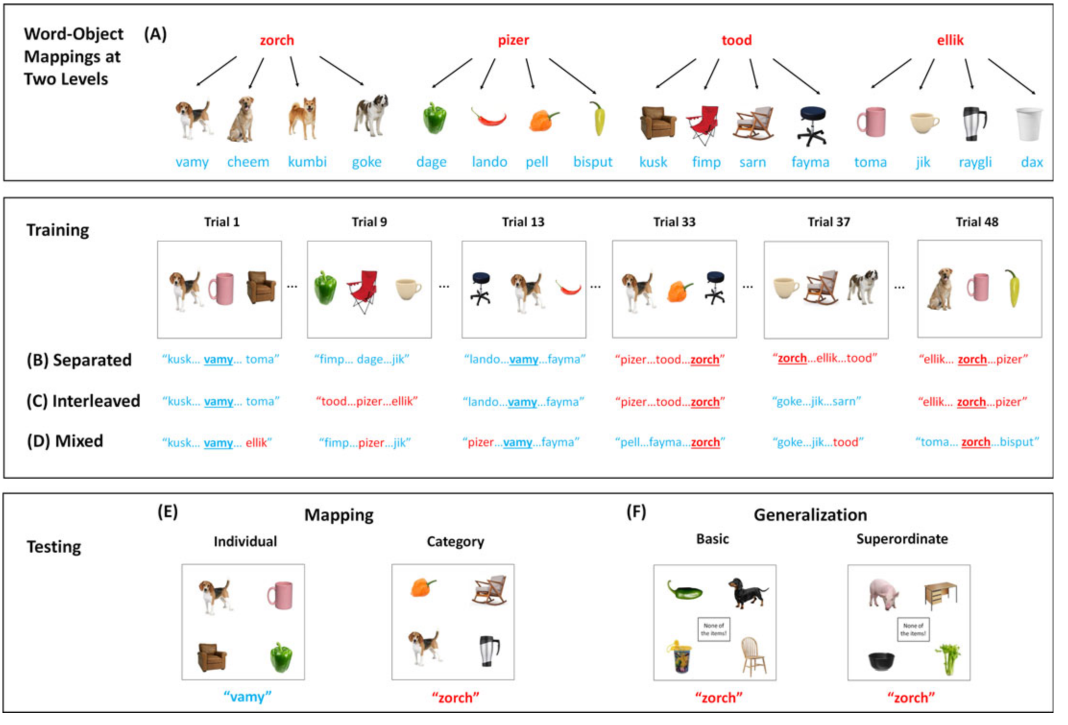

Fig. 2. Design and stimuli. (A) Training objects belonged to different basic-level categories. Each object was mapped to two labels, an individual label and a category label. The words in red represent category labels while the words in blue represent individual labels. Training (B)–(D): In each training trial, participants saw three objects and heard three words presented in random order. Participants needed to use word-object co-occurrences to learn correct word-object mappings. (B) In the Separated condition, all individual label learning trials were presented before category label learning trials. (C) In the Interleaved condition, individual label learning trials were interleaved with category label trials. (D) In the Mixed condition, each trial contained two individual labels and one category label. Testing (E) and (F): In each testing trial, participants heard one label and had to pick its referent from several options. (E) In the Individual Label Mapping trials, participants had to pick a referent from four training objects after hearing one individual label. In the Category Label Mapping trials, participants had to pick a referent from four training objects after hearing one category label. (F) In each Basic-level Generalization trial, participants heard one category label and had to select its referent from four novel objects, each from a trained basic-level category. Participants were instructed to select "none of the items" if they thought that none of the objects was a good match. In each Superordinate-level Generalization trial, participants heard one category label and had to choose from four novel superordinate-level-matched objects and a "none of the items" option.

training trials (16 objects × (6 individual labels + 3 category labels)/3 objects per trial). Each learning trial took 9 s. The entire training session lasted 7.2 min.

There were three different conditions: Separated, Interleaved, and Mixed. In the Separated condition (Fig. 2B), trials containing individual labels were presented in the first block (Individual Label Block) while trials containing category labels were presented in the second block (Category Label Block). There were 32 trials in the Individual Label Block and

16 trials in the Category Label Block. In the Interleaved condition (Fig. 2C), the training trials were divided into eight blocks, with Blocks 1, 3, 5, 7 containing individual labels and Blocks 2, 4, 6, 8 containing category labels. There were eight trials in each Individual Label Block and four trials in each Category Label Block. Each training trial in the Mixed condition contained one category label and two individual labels (Fig. 2D).

2.2.3. Testing session

Following training, there were four tasks in the testing session: Individual Label Mapping, Category Label Mapping, Basic-level Generalization, and Superordinate-level Generalization. There were 16 Individual Label Mapping and 16 Category Label Mapping trials (one for each object, see Fig. 2E). In each Mapping trial, participants heard either one individual label or one category label and had to pick its referent from four training objects, one from each category.

In the Basic-level Generalization task, participants heard one category label and saw four novel objects (Fig. 2F). Each of the four objects was a novel instance from one of the trained categories (e.g., a dachshund from the dog category). To test whether participants mapped the category label only to the instances seen during the learning session or whether they were willing to generalize the label to novel instances, each Generalization trial contained a box indicating "none of the items." The participants were instructed to find the object that best matched the label they heard in each trial. If they thought that none of the objects was a good match, they could select "none of the items." In the Superordinate-level Generalization task, participants heard one category label in each trial and saw four novel superordinate-level-matched objects (e.g., a pig). Like the Basic-level Generalization trials, each Superordinate-level Generalization trial also contained a box indicating "none of the items," which offered the option of not generalizing the category label to any of the items. This task was to examine whether participants generalized the category labels to objects outside of the training categories. Including the option "none of the items" in the Basic-level and Superordinate-level Generalization tasks allowed us to determine how willing participants were to generalize category labels to novel instances at different hierarchical levels.

2.3. Procedure

The experiment was divided into a training session and a testing session. In the training session, participants went through 48 trials, each containing three objects and three words. They had to track the word-object co-occurrences in order to learn correct mappings. Learners were only instructed to find the mappings between words and objects and were not informed of the category structure, nor were they told that there were different types of labels to learn.

Following the training session, participants first went through two Mapping tasks, Individual Label Mapping and Category Label Mapping, and then two Generalization tasks, Basic-level Generalization and Superordinate-level Generalization tasks. Each Mapping task contained 16 trials (one for each object) and each Generalization task consisted of eight trials (two for each category). These four tasks were presented in fixed order. The test trials within each task were presented in random order.

3. Results

In what follows, we first examine whether participants learned individual and category labels in each condition and compare word learning performance across conditions. After that, we investigate whether learners generalized the category labels to novel instances at different hierarchical levels and test whether there were group differences. All following analyses were conducted using mixed-effects logistic regressions (Jaeger, 2008). Response for each trial was coded as correct (1) or incorrect (0). Participants and Trials were included as random effects.

3.1. Mapping performance

We first examined whether participants were able to learn individual label-object mappings by comparing their word learning performance against chance. As shown in Fig. 3A, participants in all three conditions learned more individual labels than expected by chance (Separated: $\beta = 1.69$, $z = 6.29$, $p < .001$; Interleaved: $\beta = 1.61$, $z = 6.36$, $p < .001$; Mixed: $\beta = 0.92$, $z = 5.57$, $p < .001$). We next tested whether participants in different conditions performed differently from each other. There was a significant effect of condition ($\beta = 2.15$, $z = 6.85$, $p < .001$). Participants in both Separated and Interleaved conditions had better performance than their counterparts in the Mixed condition

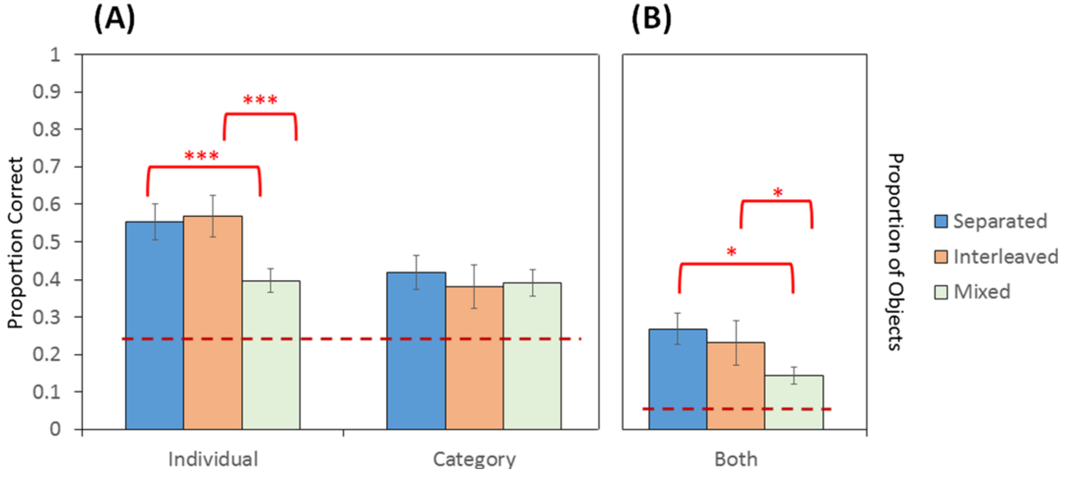

Fig. 3. Word learning performance. (A) Mean proportion (and standard error) of accurate responses in Individual and Category Label Mapping trials. The dashed line indicates chance level (0.25). (B) Mean proportion of objects that participants learned both individual and category labels for. The dashed line indicates chance level ($0.25 \times 0.25 = 0.0625$).

(βs > 2.0, ps < .001). This result indicates an advantage of blocked designs for learning individual labels.

Participants also learned more category labels than expected by chance (Separated: $\beta = 0.99$, $z = 4.42$, $p < .001$; Interleaved: $\beta = 1.19$, $z = 3.48$, $p < .001$; Mixed: $\beta = 0.89$, $z = 5.11$, $p < .001$). There was no significant group difference in their category label mapping performance ($\beta = -0.06$, $z = -0.9$, *n.s.*). This finding suggests that temporal arrangements did not affect category label learning.

As Fig. 3A shows, participants' (particularly the Separated and Interleaved groups') overall performance in the Individual Label Mapping task was better than their Category Label learning performance. These differences were not surprising, since each object co-occurred with its individual label six times but only with the category label three times during training. The accuracy differences in these two Mapping tasks likely reflected the frequency patterns in the training design.

One critical question is whether participants can learn *both* individual and category labels for the same object (Fig. 3B). All three groups of participants learned both labels for more objects than expected by chance (chance = $0.25 \times 0.25 = 0.0625$; Separated: $\beta = 1.52$, $z = 5.76$, $p < .001$; Interleaved: $\beta = 1.37$, $z = 4.24$, $p < .001$; Mixed: $\beta = 0.83$, $z = 3.70$, $p < .001$). There was a significant effect of condition ($\beta = -0.41$, $z = -2.67$, $p < .05$). Participants in both Separated and Interleaved conditions had better performance than their counterparts in the Mixed condition (βs > 0.41, ps < .05). This pattern is most likely driven by the fact that individual label learning in the Mixed condition was not as good as the other two conditions.

Together, the results from the Mapping tasks suggest that participants were able to use co-occurrences to learn both individual and category labels for the same object. As long as these two types of labels were presented in different trials, whether or not they were in completely separated blocks did not affect learning. However, individual label learning was compromised when learning trials contained both types of labels. In contrast, category label learning was not affected by temporal arrangements of label presentations.

3.2. *Generalization performance*

We next examined how broadly participants generalized the category labels. In particular, we were interested in their selection of novel same-category instances (Fig. 4A) and "none of the items" (Fig. 4B) in the Generalization tasks. If they learned the category structure in the training stimuli, they should pick the novel same-category instances in the Basic-level Generalization task but select "none of the items" in the Superordinate-level Generalization task. As Fig. 4A shows, participants in all three conditions generalized the category labels to novel same-category instances more than expected by chance in the Basic-level Generalization task (Separated: $\beta = 1.36$, $z = 4.61$, $p < .001$; Interleaved: $\beta = 0.93$, $z = 2.31$, $p < .05$; Mixed: $\beta = 0.79$, $z = 3.14$, $p < .01$) while their selections of the novel same-category instances were not different from chance in the Superordinate-level Generalization trials (all ps > .05). In contrast, they tended to select "none of the items" in the Superordinate-level Generalization task (Fig. 4B, Separated:

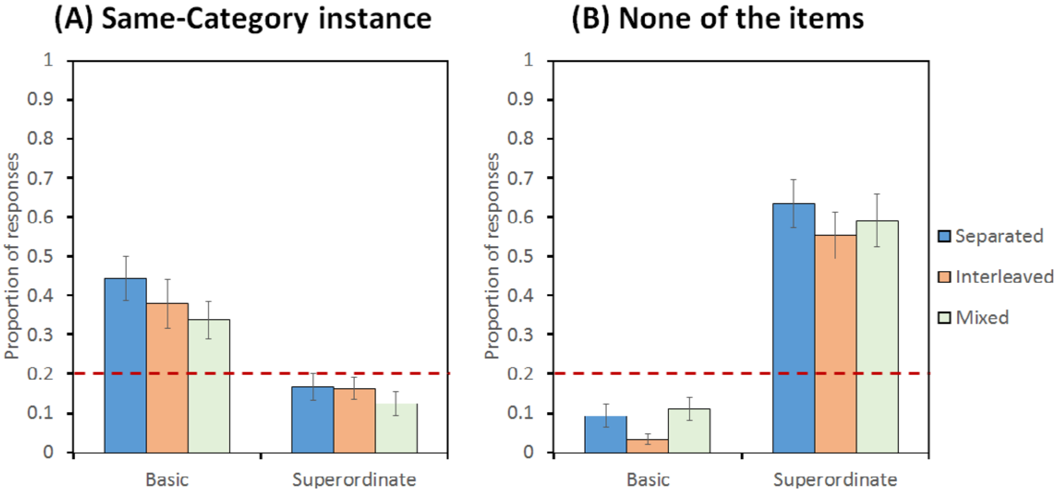

Fig. 4. Generalization performance. (A) Mean proportion (and standard error) of selecting novel same-category instances in the Basic-level and Superordinate-level Generalization tasks. There was no significant group difference within each generalization task. (B) Mean proportion of selecting "none of the items" in the Generalization tasks. The dashed lines indicate chance level (1 out of 5 possible answers, including "none of the items").

$\beta = 2.45$, $z = 6.51$, $p < .001$; Interleaved: $\beta = 1.87$, $z = 6.01$, $p < .001$; Mixed: $\beta = 2.24$, $z = 4.06$, $p < .001$), but not in the Basic-level Generalization task (Separated: $\beta = -1.462$, $z = -2.43$, $p < .05$; Interleaved: $\beta = -2.065$, $z = -2.65$, $p < .001$; Mixed: $\beta = -0.91$, $z = -2.20$, $p < .05$). There was no significant group difference in any of the response types in the generalization tasks (all ps $> .05$). These results suggest that participants in all three conditions were willing to generalize category labels to novel basic-level members, but not to superordinate-level-matched items.

4. Discussion

In this research, we investigated adults' cross-situational learning of hierarchical labels with different temporal presentation schedules. There are three main findings. First, learners are able to use cross-situational statistics to learn both individual and category labels for the same object. Second, they use co-occurrence statistics of multiple word-object mappings to infer category structure from the stimuli. Third, temporal arrangements of label presentations have different effects on the learning of individual and category labels.

4.1. Learning labels at different hierarchical levels

Previous studies have shown that adults are able to use co-occurrences to learn more than one-to-one mappings (Benitez et al., 2016; Kachergis et al., 2012; Poepsel

& Weiss, 2014; Yurovsky et al., 2013). The current research further demonstrates that participants were not only able to learn both individual and category labels when the labels were presented in different blocks. They could also learn these two types of labels at the same time in a context where labels at different hierarchical levels were presented together in the same trial. Our study suggests that adults have a very powerful statistical learning ability that allows them to concurrently track multiple levels of word-object mappings and further extract hierarchical information from the stimuli. Our results also add to the current literature by showing that humans are adept at simultaneously tracking multiple levels of statistical regularities (e.g., Chen et al., 2016; Fiser & Aslin, 2001; Monaghan, Mattock, & Walker, 2012; Romberg & Saffran, 2013).

In the current study, every Individual Label Mapping trial contained four objects, each coming from a different category. This design did not test whether learners mapped an individual label (e.g., *vamy*) to other members from the same category (e.g., other dogs). Even though we cannot rule out this possibility, we think it is unlikely in the current paradigm for three reasons. First, previous cross-situational studies using designs similar to the current one demonstrate that learners consistently map individual labels to individual target objects, but not to other members in the same category. (Chen & Yu, 2017; Chen et al., 2016). For example, Chen and Yu (2017) examined the effects of learning and retrieval contexts on cross-situational word learning by using real objects belonging to different categories (e.g., mammals, vegetables, vehicles). In Experiment 2 of that paper, there were two types of four-alternative forced-choice (4-AFC) test trials. One type of test trials had all four objects from the same category (e.g., all mammals) while the other type had each object coming from a different category. The presence of same-category distractors (i.e., foil objects from the same category as the target) did not affect test accuracy, even when category information was highlighted during both training *and* test phases. Participants mapped each label to one target object only, and not to other members from the same category. Second, as mentioned in the Method section, each training trial contained three objects, one from a different category. Therefore, each individual label never co-occurred with non-target items from the same category during learning (e.g., the word *vamy* never co-occurred with dogs other than the target beagle during training). Past cross-situational learning studies suggest that the more frequently a foil object co-occurs with a word during learning phase, the more likely it competes with the target (e.g., Roembke & McMurray, 2016). Because of the non-co-occurrence between an individual label and non-target members from the same category, those non-target members would not likely to be strong competitors even if they were included in the test trials. Third, prior word generalization research (e.g., Xu & Tenenbaum, 2007; though see also Spencer et al., 2011) suggests that, after hearing a novel word applied to one breed of dogs (e.g., beagles) multiple times, adults tend to map the word to that specific breed only and not to generalize the word to other breeds (e.g., Dalmatians, Labradors). For these reasons, we think it is unlikely that participants would map an individual label to all members of a training category in the current design.

4.2. Category label generalization

The fact that learners generalized category labels to novel instances at the same basic level suggests that they likely mapped these labels to a whole category of objects, be they present or unseen during the learning session, and not just to the trained instances. More importantly, participants' scope of generalization was consistent with the category boundaries present in the input. They generalized the labels to novel instances that came from the same basic-level categories as the trained items, but they did not over-generalize the labels to novel superordinate-level-matched instances. These results are consistent with Xu and Tenenbaum's (2007) findings and show that adults are sensitive to the input category structure.

There are a few specific features of the training categories used in the current study that are worth mentioning. The categories are basic-level categories, and thus members of a category are perceptually (fairly) similar. In addition, all training objects occurred equally often during learning phase. Previous studies have shown that the taxonomic distance between exemplars and the distributions of exemplars affect category learning and subsequent label generalization. For example, Dautriche, Chemla, and Christophe (2016, see also Dautriche & Chemla, 2016) tested adults' and 5-year-old children's category label generalization using two types of distributions, a uniform distribution and a bimodal distribution. In the uniform condition, a novel category label was mapped to a snake, a bird, a monkey, and a squirrel. In the bimodal condition, the label was mapped to two snakes and two monkeys. Both adults and children were more likely to generalize the category label to another animal in the uniform condition than in the bimodal condition. Learners took the label in the uniform condition as a word that covered a wide range of animals. In contrast, they took the label in the bimodal condition as a homophone that only covered the snake and monkey categories. A few potentially fruitful directions for further investigation concern the roles of taxonomic or perceptual distance between exemplars, exemplar distributions, and size of categories in category label learning and word generalization.

Another characteristic of the present study is that the objects and categories are all familiar to adult participants, which is a design used in many previous word generalization studies (e.g., Spencer et al., 2011; Xu & Tenenbaum, 2007). Jenkins, Samuelson, Smith, and Spencer (2015) found that how broadly 3- to 5-year-old children generalized a novel category label was not only affected by the exemplars they saw when they heard the label, but also by their prior category knowledge. Therefore, another area for future research is how prior knowledge about objects or categories affects category label learning and generalization and how these processes change across development.

4.3. Effect of temporal arrangements on word learning

The current research also shows that different temporal presentations only affected the learning of individual labels but not category labels. This result indicates that temporal arrangements differentially affected the learning of different types of labels. It was easier to learn individual labels when they were presented in blocks, be they completely separated from or interleaved with blocks containing category labels. The presence of a category label in the same trial significantly interfered with the learning of individual labels. The benefit of blocked designs in learning individual labels is consistent with previous cross-situational learning studies (e.g., Benitez et al., 2016; Vlach & Johnson, 2013; Yurovsky et al., 2013). Yet, unlike previous research showing a positive interleaved effect in category learning (e.g., Kang & Pashler, 2012; Kornell & Bjork, 2008; Wahlheim et al., 2011), our study suggests that there was no advantage or disadvantage of presenting category labels in separated, interleaved, or mixed schedules.

Why is individual label learning affected by temporal arrangement of stimuli? One possibility is that different word-object mapping principles may be applied in different types of learning trials when learners aggregate statistical information across trials. In blocked designs, all items in an individual label trial had the same mapping principle (i.e., mapping to one, and only one, object across trials). However, in a mixed design, labels in a trial followed different mapping principles (i.e., two labels were mapped to only one object across trials and one label was mapped to multiple objects across trials). The inconsistency in the mapping principles for different labels in the same trial likely made it hard to track word-object co-occurrences across trials. This result also indicates that the additional regularity of having the same mapping principle may be beneficial for individual label learning. One topic for future work is how temporal presentation interacts with the accumulation or integration of information and how that influences the learning of different types of words.

Conclusions

The current research demonstrates that adults are able to extract multiple levels of word-object co-occurrences and use them to learn hierarchical labels. In addition, adults generalize category labels to novel instances, which is indicative of word-concept associations. Importantly, learners' scope of generalization is consistent with the category boundary present in the input, suggesting their sensitivity to the input structure. Our results also show that different types of words are differentially affected by different presentation schedules.

Acknowledgments

This work was supported by NIH R01HD074601. We thank Eric Timperman for help with data collection.

References

Akhtar, N., & Montague, L. (1999). Early lexical acquisition: The role of cross-situational learning. *First Language*, *19*(57), 347–358.

Baldwin, D. A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology*, *29*(5), 832–843.

Benitez, V. L., Yurovsky, D., & Smith, L. B. (2016). Competition between multiple words for a referent in cross-situational word learning. *Journal of Memory and Language*, *90*, 31–48.

Carvalho, P. F., & Goldstone, R. L. (2015). The benefits of interleaved and blocked study: Different tasks benefit from different schedules of study. *Psychonomic Bulletin & Review*, *22*(1), 281–288.

Chen, C., Gershkoff-Stowe, L., Wu, C. Y., Cheung, H., & Yu, C. (2016). Tracking multiple statistics: Simultaneous learning of object names and categories in English and Mandarin speakers. *Cognitive Science*. https://doi.org/10.1111/cogs.12417

Chen, C., & Yu, C. (2017). Grounding statistical learning in context: The effects of learning and retrieval contexts on cross-situational word learning. *Psychonomic Bulletin & Review*, *24*(3), 920–926. Advance online publication. https://doi.org/10.3758/s13423-016-1163-x.

Childers, J. B., & Paik, J. H. (2009). Korean-and English-speaking children use cross-situational information to learn novel predicate terms. *Journal of Child Language*, *36*, 201–224.

Clark, E. V., & Svaib, T. A. (1997). Speaker perspective and reference in young children. *First Language*, *17* (49), 57–74.

Dautriche, I., & Chemla, E. (2016). What Homophones Say about Words. *PLoS ONE*, *11*(9), e0162176.

Dautriche, I., Chemla, E., & Christophe, A. (2016). Word learning: Homophony and the distribution of learning exemplars. *Language Learning and Development*, *12*(3), 231–251.

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, *12*(6), 499–504.

Gangwani, T., Kachergis, G., & Yu, C. (2010). Simultaneous cross-situational learning of category and object names. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Conference of the Cognitive Science Society* (pp. 1595–1600). Austin, TX: Cognitive Science Society.

Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, *28*(1), 99–108.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.

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.

Kachergis, G., Yu, C., & Shiffrin, R. M. (2012). An associative model of adaptive inference for learning word–referent mappings. *Psychonomic Bulletin & Review*, *19*(2), 317–324.

Kang, S. H., & Pashler, H. (2012). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology*, *26*(1), 97–103.

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories is spacing the "enemy of induction"? *Psychological Science*, *19*(6), 585–592.

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

Monaghan, P., Mattock, K., Davies, R. A., & Smith, A. C. (2015). Gavagai is as Gavagai does: Learning nouns and verbs from cross-situational statistics. *Cognitive Science*, *39*(5), 1099–1112.

Monaghan, P., Mattock, K., & Walker, P. (2012). The role of sound symbolism in language learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(5), 1152–1164.

Poepsel, T. J., & Weiss, D. J. (2014). Context influences conscious appraisal of cross-situational statistical learning. *Frontiers in Psychology*, *5*, https://doi.org/10.3389/fpsyg.2014.00691.

Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

Roembke, T. C., & McMurray, B. (2016). Observational word learning: Beyond propose-but-verify and associative bean counting. *Journal of Memory and Language*, *87*, 105–127.

Romberg, A. R., & Saffran, J. R. (2013). All together now: Concurrent learning of multiple structures in an artificial language. *Cognitive Science*, *37*, 1290–1318.

Scott, R. M., & Fisher, C. (2012). 2.5-year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition*, *122*(2), 163–180.

Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism? *Cognition*, *60*(2), 143–171.

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

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.

Suanda, S. H., & Namy, L. L. (2012). Detailed behavioral analysis as a window into cross-situational word learning. *Cognitive Science*, *36*, 545–559.

Tomasello, M., & Akhtar, N. (1995). Two-year-olds use pragmatic cues to differentiate reference to objects and actions. *Cognitive Development*, *10*(2), 201–224.

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition*, *127*(3), 375–382.

Vlach, H. A., & Kalish, C. W. (2014). Temporal dynamics of categorization: Forgetting as the basis of abstraction and generalization. *Frontiers in Psychology*, *5*. doi:10.3389/fpsyg.2014.01021

Wahlheim, C. N., Dunlosky, J., & Jacoby, L. L. (2011). Spacing enhances the learning of natural concepts: An investigation of mechanisms, metacognition, and aging. *Memory & Cognition*, *39*(5), 750–763.

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., Yu, C., & Smith, L. B. (2013). Competitive processes in cross-situational word learning. *Cognitive Science*, *37*(5), 891–921.