Journal of Experimental Child Psychology 200 (2020) 104961

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

The temporal structure of naming events differentially affects children's and adults' cross-situational word learning

Viridiana L. Benitez [a,*], Martin Zettersten [b], Erica Wojcik [c]

[a] Department of Psychology, Arizona State University, Tempe, AZ 85287, USA
[b] University of Wisconsin–Madison, Madison, WI 53706, USA
[c] Department of Psychology, Skidmore College, Saratoga Springs, NY 12866, USA

ARTICLE INFO

Article history:
Received 5 February 2020
Revised 7 May 2020

Keywords:
Cross-situational learning
Statistical learning
Word learning
Children
Adults
Language development

ABSTRACT

To acquire novel words, learners often need to integrate information about word meanings across ambiguous learning events distributed in time. How does the temporal structure of those word learning events affect what learners encode? How do the effects of temporal structure differ in children and adults? In the current experiments, we asked how 4- to 7-year-old children's (N = 110) and adults' (N = 90) performance on a cross-situational word learning task is influenced by the temporal distribution of learning events. We tested participants in three training conditions, manipulating the number of trials that separated naming events for specific objects. In the Unstructured condition, the temporal distribution was varied; in the Massed condition, naming events occurred with few interleaved trials; and in the Interleaved condition, naming events occurred with many interleaved trials. Adults showed substantially larger benefits from the Massed condition than children, whereas children were equally successful at learning in the Massed and Interleaved conditions. These results provide evidence that adults differ from children in how they exploit temporal structure during cross-situational word learning.

* Corresponding author.
E-mail address: viridiana.benitez@asu.edu (V.L. Benitez).

https://doi.org/10.1016/j.jecp.2020.104961

Introduction

Word learning is a process that unfolds *over time* across learning events that are often *ambiguous.* Forming a robust representation of a word is a temporally distributed learning process that relies on repeated exposure. Whereas the first exposure to a new word can lead to the rapid encoding of a word–referent association (Carey & Bartlett, 1978; Halberda, 2003), building a strong lexical–semantic representation requires reinforcing and enriching meaning across numerous learning events (Horst, 2013; Kucker, McMurray, & Samuelson, 2015). Furthermore, word learning events are often ambiguous. Children may hear words without an ostensive cue or clear context to suggest the exact referent for the label (Medina, Snedeker, Trueswell, & Gleitman, 2011; Quine, 1960). Thus, the key to successful word learning is resolving referential ambiguity across multiple word learning events that are distributed in time. But how does the temporal distribution of those learning events affect what words are learned? Although we know much about how word learners resolve referential ambiguity broadly, we know less about how the distribution of ambiguous naming events in time influences learning the correct word–referent links across development (Smith, Suanda, & Yu, 2014; Yu & Smith, 2012; Yurovsky & Frank, 2015).

Resolving referential ambiguity in word learning

Many individual naming events are referentially ambiguous (Quine, 1960). However, referential ambiguity can be resolved across several learning moments rather than within one learning moment (Yu & Smith, 2012). A host of experimental evidence shows that learners can overcome referential uncertainty by tracking information across multiple learning events, a process termed *cross-situational word learning* (see Yu & Smith, 2012, for a review). In one common paradigm first developed by Yu and Smith (2007), learners are presented with a series of object naming trials. On each trial, several unfamiliar objects are presented with an array of novel labels. Although it is not clear from any single trial which label refers to which object, across trials the probability of a specific referent co-occurring with a specific label is 1.0 and the probability of other referents co-occurring with that label is far lower. After training, infants as young as 12 months (Smith & Yu, 2008), children (Suanda, Mugwanya, & Namy, 2014), and adults (Yu & Smith, 2007) perform significantly better than chance at identifying the correct referents for the novel labels. Both behavioral data and computational models of performance in this task have provided evidence that learners are able to acquire words in ambiguous contexts by aggregating co-occurrence statistics over time (Blythe, Smith, & Smith, 2010; Ichinco, Frank, & Saxe, 2009; Kachergis, Yu, & Shiffrin, 2012; Regier, 2005; Siskind, 1996; Yu & Ballard, 2007; Yu & Smith, 2012; but see Trueswell, Medina, Hafri, & Gleitman, 2013, for an alternative hypothesis).

The temporal structure of naming events

In addition to referential ambiguity, a second fundamental characteristic of word learning events is that they unfold in systematic patterns across time. Recent naturalistic investigations have begun to document the nature of this temporal structure, identifying features of naming events that may aid learners when encountering novel words (Smith, Jayaraman, Clerkin, & Yu, 2018). For example, caregiver naming events often exhibit a "bursty" (massed) structure, in which a word will repeat across multiple conversational turns but then not again for several hours or days. During caregiver–child play with objects, caregiver discourse that includes consecutive repetitions of object labels has been linked with high-quality word learning moments as well as with children's vocabulary knowledge (Frank, Tenenbaum, & Fernald, 2013; Schwab, Rowe, Cabrera, & Lew-Williams, 2018; Suanda, Smith, & Yu, 2016). These studies suggest that in naturalistic language experiences, the distribution of naming events across time exhibits structure that children can exploit in learning new words.

However, although naturalistic observations have begun to characterize the temporal structure of children's exposure to novel words, we still know little about how this temporal structure affects word learning, particularly when children are confronted with multiple ambiguous learning moments across time. This is surprising because aggregating word–object co-occurrences across ambiguous situations is a process that, by definition, has a temporal structure. Word learning across multiple events requires learners to register word–referent co-occurrences presented in a given instance, maintain this information over time, and retrieve it when the word occurs again. The temporal distribution of naming events therefore may be critical for correctly tracking word–referent co-occurrences. That is, whether an object and its name are repeated close in time (massed), versus further apart in time with other objects named in between (interleaved), may affect how learners strengthen each word–referent link.

There are several reasons why temporal structure should be particularly important for understanding cross-situational word learning. First, the temporal distribution of events affects learning and memory processes, with massed events boosting immediate encoding and interleaved events boosting long-term retention (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006). Second, proposals about statistical learning in general and cross-situational word learning in particular have emphasized the importance of attention and memory constraints, both processes that are modulated by temporal structure (Goldstein et al., 2010; Yu & Smith, 2012). Third, if we are to understand the mechanisms underlying ambiguous word learning over time, we must begin to incorporate features of naturalistic language experiences in our studies examining cross-situational word learning. Therefore, it is necessary to consider how the temporal distribution of ambiguous naming events over time affects word–referent learning.

Temporal structure is known to influence word learning in unambiguous, ostensive naming events. Past work has shown that whether ostensive learning events have a massed versus interleaved structure affects young children's encoding and retention of new words (Childers & Tomasello, 2002; Schwab & Lew-Williams, 2016; Vlach, Ankowski, & Sandhofer, 2012). In particular, when object naming events have massed structure, in which the label for a given object is repeated in succession, young children encode object names better compared with naming events that interleave the labeling of different objects during training (Schwab & Lew-Williams, 2016).

Although laboratory studies examining ostensive word learning suggest that massed structure may be beneficial for word learning, it is unclear how these results should generalize to ambiguous naming events. Unlike in ostensive word learning tasks, repeated naming events in cross-situational word learning are not identical. Although the learner will both hear the target label and see the target object on each repetition, each trial also includes a rotating cast of distractor labels and objects. Thus, whereas massed naming events in ostensive word learning present only one word–referent association to be encoded, massed naming events in ambiguous word learning present several. Although a massed structure may still benefit the encoding of the correct word–referent pair in this situation, it is possible that the increased complexity of the information within and across trials could alter the effect of temporal structure on learning.

Temporal structure in cross-situational word learning

A few studies have examined the effects of temporal structure on cross-situational word learning in adults and children, albeit separately. To examine the mechanisms underlying cross-situational word learning, two studies testing adults manipulated a number of different training factors, including the temporal structure of repetitions of words (Smith, Smith, & Blythe, 2011; Yurovsky & Frank, 2015). During the training phase of both studies, adults were presented with one target word and asked to select its target object out of multiple objects presented on each trial. In Smith et al. (2011), target words occurred according to either massed (in consecutive trials) or interleaved (with several trials in between repetitions) training schedules, whereas Yurovsky and Frank (2015) included several conditions that manipulated the number of trials interleaved between word repetitions. Their results showed that adults were more likely to select the correct target object when repetitions of a word occurred consecutively (Smith et al., 2011) or with few interleaved trials (Yurovsky & Frank, 2015). Consistent with these findings, Kachergis, Shiffrin, and Yu (2009) demonstrated that adults' accuracy improved when consecutive repetitions of word–object pairings were allowed to occur in a cross-situational word learning task presenting four words and four objects on each training trial (see also

Onnis, Edelmann, & Waterfall, 2011). Furthermore, adults' accuracy increased as a function of the number of word–object pairings that repeated in consecutive trials. Together, these findings suggest that for adults, massed naming structure provides a benefit to cross-situational word learning.

In contrast, past research suggests that children tend to perform similarly for massed and interleaved structure during cross-situational word learning (Smith & Yu, 2013; Vlach & Johnson, 2013). In these studies, children were presented with two words and two objects on each training trial. Training trials were structured such that on each trial, one of the two objects repeated (massed objects), whereas the second object always changed (interleaved objects). Therefore, temporal structure was a within-participant condition; although all words were learnable via cross-situational statistics, the presentation of a given object was in either a massed or interleaved format. When testing 12- and 14-month-olds, infants as a group did not show evidence of learning (Smith & Yu, 2013). However, in a study with older infants, 16-month-olds learned the massed objects but not the interleaved objects, whereas 20-month-olds learned both. Finally, Vlach and DeBrock (2017) presented a similar task to preschoolers to examine individual differences in cross-situational word learning. Their results showed that preschoolers were accurate for both massed and interleaved objects. Together, these findings indicate that as children get older, they are successful with both massed and interleaved structure during cross-situational word learning.

Although this research has started to shed light on the effects of temporal structure on cross-situational word learning, several important questions remain. The first question concerns disentangling the independent effects of massed versus interleaved structure on the time course of cross-situational word learning. In particular, in the previous cross-situational word learning studies examining temporal structure, items trained via massed or interleaved structure were included within the same learning task. If there are differential effects of massed structure compared with interleaved structure on word learning, including items trained with each type of structure in the same task is likely to lead to interaction effects. For example, in Vlach and DeBrock (2017), children saw both massed and interleaved word–object pairings on each training trial. If massed structure provides a benefit for word learning, then massed objects may have helped children to disambiguate interleaved objects in this task, masking the differential impact of each structure on word learning. Understanding the influence of temporal structures on cross-situational word learning requires assessing how massed structure may affect ambiguous word learning independent of interleaved structure and vice versa.

A second unresolved question is whether there are developmental differences with respect to how temporal structure affects cross-situational word learning. Previous research suggests that there may be systematic differences in the degree to which child and adult learners benefit from temporal structure. Specifically, older children are successful across massed and interleaved temporal structures (Vlach & DeBrock, 2017), whereas adults are more accurate when naming events for a given object are massed rather than interleaved (Onnis, Edelmann, & Waterfall, 2011; Smith, Smith, & Blythe, 2011; Yurovsky & Frank, 2015; Kachergis et al., 2009). However, the testing conditions for studies with adults differ in numerous ways from those with children, making it difficult to directly compare children's and adults' performance. For example, children have been tested in designs with fewer words and objects presented on a given trial than adults. It is possible that temporal structure affects cross-situational word learning only under more complex conditions such as those used in studies with adults. Understanding how the effects of temporal structure on cross-situational word learning change with age requires testing learners under the same conditions, something that has rarely been done in previous cross-situational research (Bunce & Scott, 2017; Fitneva & Christiansen, 2017).

Beyond addressing the methodological limitations of previous research, direct comparisons of children's and adults' performance on cross-situational word learning is important for understanding the underlying mechanisms behind ambiguous word learning. Although many proposals exist to explain how learners approach cross-situational word learning (Berens, Horst, & Bird, 2018; Kachergis et al., 2012; Kachergis, Yu, & Shiffrin, 2017; Smith et al., 2011; Stevens, Gleitman, Trueswell, & Yang, 2017; Trueswell et al., 2013; Yu & Smith, 2012; Yurovsky & Frank, 2015), all of these proposals have been developed based mainly on data from adult learners. However, key processes that have been implicated in cross-situational word learning, including memory, attention, and learning strategies, change dramatically over development. In theoretical debates surrounding how *children* learn words, cross-situational word learning is one of the main hypothesized mechanisms (Kucker et al., 2015; Regier, 2005; Siskind, 1996; Smith et al., 2018). Therefore, any model of cross-situational word learning that does not use data from children is incomplete. Directly comparing adult and child cross-situational word learning allows researchers to determine whether the existing models based on adult data are applicable to child language learning and what differences in learning performance must be considered.

The current study

In this study, we set out to compare how temporal structure affects cross-situational word learning in children and adults by (a) separately examining how massed structure and interleaved structure affect learning during a cross-situational word learning task and (b) testing children and adults under identical conditions. We tested 4- to 7-year-old children and adults in the same cross-situational word learning task while manipulating the temporal structure of naming events between participants. On any individual trial, learners were presented with 2 words and 2 objects, with no clear indication as to which words went with which objects. Across all trials, however, each word co-occurred most often with a single target object.

All learners were presented with the same task with one key manipulation: Across three between-participant conditions, we changed the temporal structure of naming events such that there were differences in the number of trials in between repeated naming events for a given object. In the Unstructured condition, training trials were randomly shuffled, creating a mix of massed and interleaved naming events, similar to previous studies. The number of trials in between naming events for a given object ranged from 1 to 6 (mean number of trials = 3). Unlike previous studies, however, we included two additional conditions with different distributions of learning events across time: Massed and Interleaved. In the Massed condition, naming events were structured such that in between object repetitions, few trials occurred naming other objects (range = 0–2 trials, M = 0.75). In contrast, in the Interleaved condition, many more trials occurred in between naming events for a given object (range = 2–4 trials, M = 3).

In Experiment 1, children were presented with the Unstructured condition. Experiment 2 was a preregistered study testing two new groups of children in either the Interleaved condition or the Massed condition (https://osf.io/g2md4). In preregistered Experiment 3, adults were presented with one of the three temporal structure conditions (https://osf.io/84trw). All learners were immediately tested on their knowledge of the correct word–referent mappings. The key questions were (a) how learning is affected by the temporal structure of naming events and (b) whether children and adults differ in how they exploit temporal structure. To examine the first question, we compared learners' performance across all three structure conditions. If massed structure, independent from interleaved structure, benefits cross-situational word learning, then we should find that learners perform better in the Massed condition than in the other two conditions. Critically, we can also assess how interleaved structure, independent of massed structure, contributes to learning by comparing the Interleaved condition with the Unstructured condition, which mixes both types of structure. Finally, by comparing child and adult performance, we can also clarify whether there are age differences in how temporal structure affects cross-situational word learning. If all learners benefit from massed structure, then both children and adults should perform best in the Massed condition. If, however, developmental differences influence how learners exploit temporal structure during cross-situational word learning, then we should find a significant interaction between age group and temporal structure.

Experiment 1

In Experiment 1, we presented children with a cross-situational word learning task where each of 8 objects was labeled three times at training. On each trial, children saw 2 objects and heard 2 words. Critically, the training list was manipulated to be a random mix of naming events (Unstructured condition) such that the number of trials in between repetitions of objects varied from 1 to 6. Testing immediately followed; children were given a target word and needed to select the target object from two referents. The goal in Experiment 1 was to assess whether children could track word–referent pairings given randomly structured naming events.

Method

Participants

A total of 30 4- to 7-year-old-children (M_{age} = 5.73 years, SD = 1.01, range = 4–7.8) were recruited and tested as part of the Living Lab Initiative at the Children's Museum of Phoenix in the southwestern United States. An additional 6 children were tested but excluded due to displaying a side bias (selecting the object on the same location on 7 or more of the 8 test trials; n = 4) or for failing to complete the task (n = 2). The final sample included 18 girls, included 4 children whose parents reported they were exposed to a second language in addition to English,[1] and was predominantly White (24 children were identified by their parents as White, 1 as Hispanic, 1 as Asian, and 3 as two or more categories; 1 parent did not respond to the race question). Most of the sample's parents had a bachelor's degree or higher (n = 20), with 7 parents indicating some college or an associate's degree and 1 parent with no high school diploma (2 parents did not indicate their education). Children received a sticker for participating. Parental consent was obtained for all participants according to the Arizona State University institutional review board.

Stimuli

In total, 10 pictures of novel objects and 10 English-like novel words were selected from the Novel Objects and Unusual Names database (NOUN; Horst & Hout, 2016). Half of the words were one-syllable items, and the other half were two-syllable items. Words were recorded in isolation by a native English-speaking woman in a child-friendly voice (all stimuli used for these experiments can be found at https://osf.io/2hmxr). Each word was randomly paired with a corresponding object to create two sets of 8 word–object pairings to be trained and tested. An additional 2 word–object pairings were chosen to be presented during familiarization. Word–object pairing set was counterbalanced across children.

Design

Training. During training, children were presented with 12 trials. Each training trial presented 2 objects and 2 words; both objects were presented first, followed 1 s later by the first word (presented auditorily for 1 s), a 0.5-ms pause, and then the second word (for 1 s). Objects were on the screen for a total of 5 s. The order of object position on the screen and the sequential presentation of the words was randomized across trials. Therefore, within a given trial, there was no clear indication as to which word referred to which object. Across trials, however, each object occurred with its target word three times and occurred with 3 distractor objects one time.

Training trials were randomly intermixed with the constraint that objects could not repeat on consecutive trials (see Fig. 1–Unstructured). This resulted in a range of 1 to 6 interleaving trials between repetitions of an object during which other objects were named (mean number of trials = 3). Two training orders were created and counterbalanced across children.

Testing. Testing immediately followed training. Children were tested using a two-alternative forced-choice (2AFC) method. Children were presented auditorily with the target word while viewing the target and a distractor, and they were asked to touch the target referent. Each word–object pairing was tested once, creating a total of 8 test trials. Distractor objects were objects that had occurred once with the target word. After each selection, children were presented with a smiley face and were not provided with any feedback other than being thanked for their selection.

[1] We assessed whether second language exposure had an effect on accuracy scores. For this experiment and the following experiments, we found no effect of second language exposure on accuracy and no interaction between second language exposure and condition. Therefore, we included all participants in our final samples regardless of second language exposure.

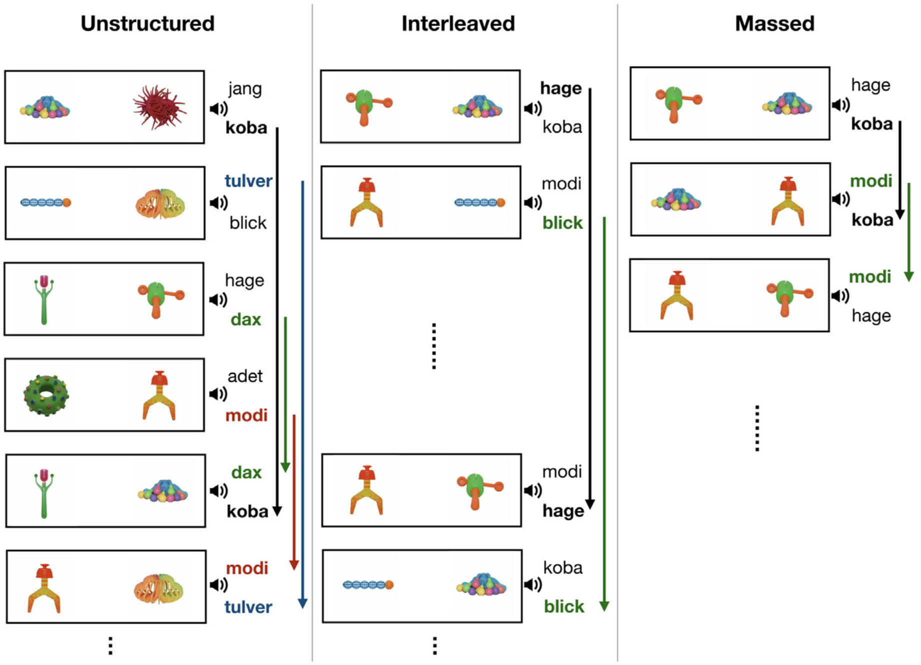

Fig. 1. Training design for the Unstructured condition (Experiments 1 and 3) and the Interleaved and Massed conditions (Experiments 2 and 3). The arrows represent the number of trials separating naming events for a specific label.

Procedure

The task was programmed in PsychoPy Builder and presented on a Dell Latitude 5285 touchscreen laptop connected to an external speaker (materials are available at: https://osf.io/2hmxr). Testing was conducted in a main area of the children's museum. Children sat in front of the laptop while an experimenter sat next to them and guided them through the task.

The task began by introducing children to Ella the elephant, who just had a birthday and was gifted with many new toys. Children were shown a picture of Ella, followed by pictures of all novel objects to be learned as well as the auditory presentation of the object names in a random order (Suanda et al., 2014). Children were instructed to try and remember all of Ella's toy names because Ella would ask them about the names later. Children were then instructed that they would practice, at which point they were presented with a familiarization phase. During the familiarization phase, children were shown two novel toys, one at a time, with the auditory presentation of each toy's corresponding name. They were then presented with two 2AFC practice test trials in which they were asked to select the target object of each of the 2 familiarized novel words. Children were provided with feedback and were corrected if they chose the incorrect object. After familiarization, children were presented with the training phase, followed by the testing phase. Parents were asked to complete a short demographic questionnaire during or after the task. The entire session lasted approximately 10 min.

Results

All data and R scripts of the analyses are openly available on the Open Science Framework (https://osf.io/2hmxr).

Relationship between age and accuracy

First, given our large age span, we examined whether age had an effect on children's task accuracy. To test this, we aggregated children's accuracy over the test trials to obtain one mean accuracy score for each child. We conducted a simple linear regression using mean accuracy as the dependent variable and age as the independent variable. Results showed no significant effect of age on accuracy, $b = -0.0006$, $t(28) = -0.02$, $p = .99$ (see Fig. 2).

Overall accuracy

Fig. 3A displays children's mean accuracy at test for Experiment 1. Our main question was whether children were able to track the word–referent co-occurrences given the random mix of training trials in our task. To test this, we compared children's performance against chance (.50) using the *lme4* package (Version 1.1–21) in R (Bates, Mächler, Bolker, & Walker, 2015; R Development Core Team, 2019) to fit a logistic mixed-effects model predicting children's trial-by-trial test accuracy, including a simple intercept [note that logit(0.5) = 0] and by-participant and by-item random intercepts. Results showed that children's accuracy ($M = .56$, $SD = .22$, 95% confidence interval (CI) = [.48, .64]) did not differ from chance performance, $b = 0.26$, Wald 95% CI = [−0.15, 0.66], Wald $\chi^2(1) = 1.57$, $p = .21$, suggesting that as a group children did not show substantial evidence of succeeding at cross-situational word learning.

Discussion

The results for Experiment 1 revealed that children were not successful as a group at learning the correct word–referent mappings. Our random mix of training trials generated a range of 1 to 6 interleaved trials between object repetitions, with an average of 3 interleaved trials. Although there are several different factors that may have yielded a lack of evidence for learning (which we come back to in the General Discussion), it is possible that children failed in part because of this random temporal structure. These findings establish that the current cross-situational learning task is somewhat difficult for children—at least in the absence of temporal structure. The goal in Experiment 2 was to investigate how structuring repetitions of naming events closer in time benefits children's learning over structuring repetitions of naming events further apart in time.

Experiment 2

In preregistered Experiment 2 (https://osf.io/g2md4), we presented two new groups of children with a similar version of the cross-situational word learning task presented in Experiment 1. However, we manipulated the training structure such that naming events were either massed, where 0 to 2 training trials were presented in between repetitions of a given object, or interleaved, where 2 to 4 trials were presented in between repetitions of a given object. Testing immediately followed. Across both conditions (Massed and Interleaved), the word–referent co-occurrences were the same.

If massing object repetitions aids children's cross-situational word learning, as suggested by previous research in ostensive word learning (Schwab & Lew-Williams, 2016), then children in the Massed condition should outperform children in the Interleaved condition. However, if children learn equally well from both massed and interleaved temporal structure, as suggested by previous research in cross-situational word learning (Vlach & DeBrock, 2017), then there should be no differences in performance between the Massed and Interleaved conditions. Furthermore, if children in general have a hard time learning from conditions in which many training trials are interleaved between repetitions of objects, as suggested by our Experiment 1 results, then we should find that children will not succeed at learning in the Interleaved condition.

Method

Participants

A new group of 80 4- to 7-year-old children (ranging in age from 4 to 7.9 years) who did not participate in Experiment 1 participated in Experiment 2. As in Experiment 1, children were recruited and tested as part of the Living Lab Initiative at the Children's Museum of Phoenix. An additional 15 children were tested but excluded due to displaying a side bias at test (Massed: $n = 5$, Interleaved: $n = 6$),

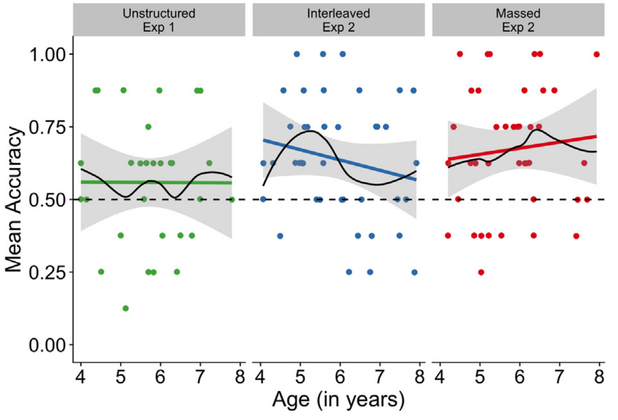

Fig. 2. Relationship between participant age and mean accuracy in Experiments 1 and 2. Colored lines and error bands represent linear fits predicting accuracy from age (no significant effects). Black lines show LOESS (local regression) fits to represent any local relationships across age. (For interpretation of the reference to color in this figure legend, the reader is referred to the Web version of this article.)

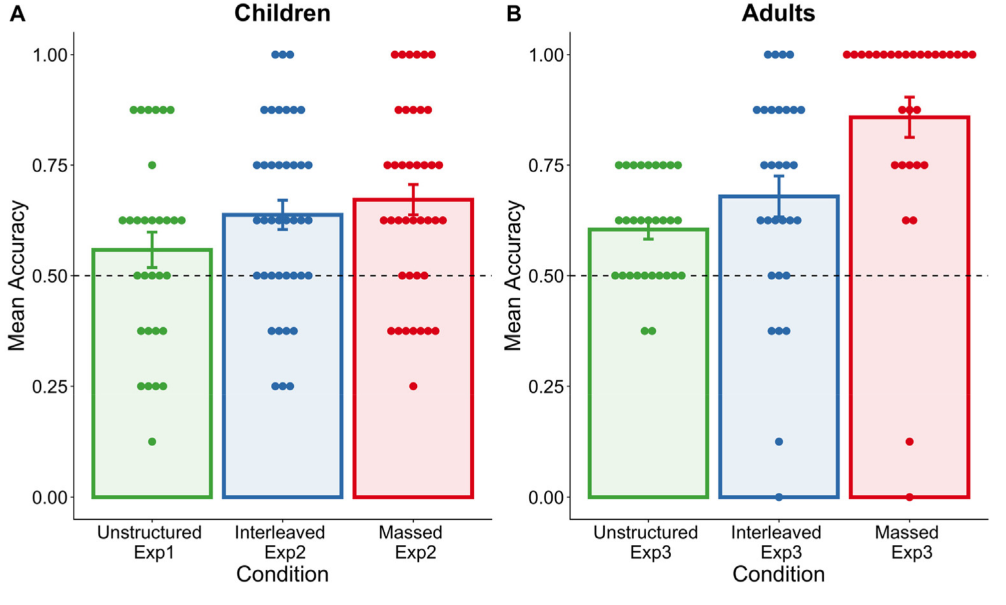

Fig. 3. Mean accuracy across condition for children (Experiments 1 and 2) (A) and adults (Experiment 3) (B). Error bars represent standard errors of the mean.

for failing to follow instructions at test (n = 1), for not completing the task (n = 1), for computer error (n = 1), and for parental report of atypical development (n = 1). The final sample included 56 girls, included 20 children whose parents reported they were exposed to a second language in addition to English, and identified predominantly as White (50 children were identified by their parents as White, 5 as African American, 9 as Hispanic, 5 as Asian, and 10 as two or more categories; 1 parent did not respond to the race question). Most of the sample's parents had a bachelor's degree or higher

(n = 51), with 20 parents indicating some college or an associate's degree, 7 indicating having a high school diploma, and 2 with no high school diploma. All children received a sticker for participating.

Children were randomly assigned to one of two conditions: the Massed condition (n = 40; M_{age} = 5.79 years, SD = 1.00) or the Interleaved condition (n = 40; M_{age} = 5.94 years, SD = 1.13). Ages were comparable between conditions, $t(78) = 0.63$, $p = .53$, and also were comparable to the ages of children in Experiment 1, $t(108) = 0.63$, $p = .53$.

Stimuli

A total of 10 new pictures of novel objects were selected from the NOUN database for Experiment 2 (Horst & Hout, 2016). The words were the same as those used in Experiment 1. As in Experiment 1, each word was randomly paired with a corresponding object to create two sets of 8 word–object pairings. Presentation of a set was counterbalanced across children.

Design

Training. Children were presented with the same task as in Experiment 1 with a key change: The order of training trials was manipulated to be either massed or interleaved. In the Massed condition, there were no trials or only a few trials in between repetitions of objects (range of interleaved trials = 0–2, mean number of interleaved trials across objects and repetitions = 0.75). In the Interleaved condition, there were many more interleaved trials in between repetitions of objects (range of interleaved trials = 2–4, mean number of interleaved trials = 3). Thus, the average number of interleaved trials was lowest in the Massed condition and highest in the Interleaved condition. The average for the Interleaved condition was the same as for the Unstructured condition in Experiment 1, although the range of interleaved trials differed across all three conditions.

Testing. Testing was exactly the same as in Experiment 1.

Procedure

The procedure was exactly the same as in Experiment 1.

Results

Relationship between age and accuracy

As in Experiment 1, in preregistered analyses, we examined whether age had an effect on children's accuracy in Experiment 2. Results showed no significant effect of age on accuracy for either the Massed condition, $b = 0.02$, $t(38) = 0.60$, $p = .55$, or the Interleaved condition, $b = -0.04$, $t(38) = -1.20$, $p = .24$, and no significant interaction between age and condition, $b = 0.06$, $t(76) = 1.24$, $p = .22$ (see Fig. 2).

Accuracy in the Interleaved versus Massed conditions

Fig. 3A displays children's mean accuracy at test for both conditions for Experiment 2. Our main question was whether children's performance varied between the two types of training conditions. To test this, in preregistered analyses, we compared children's performance across conditions by fitting a logistic mixed-effects model predicting children's trial-by-trial test accuracy from condition (Interleaved vs. Massed; dummy coded). We first included by-participant and by-item random intercepts and a by-item random slope for condition. When the initial model did not converge, we followed the recommendations of Barr, Levy, Scheepers, and Tily (2013) and simplified the random-effects structure by removing the by-item random slope for condition (results remained qualitatively the same). The fixed effect of condition did not improve the fit of the model, $b = 0.17$, 95% Wald CI = [−0.27, 0.62], Wald $\chi^2(1) = 0.57$, $p = .45$, suggesting that there were no significant differences between the Massed condition (M = .67, SD = .22, 95% CI = [.60, .74]) and the Interleaved condition (M = .64, SD = .21, 95% CI = [.57, .70]). Children were above chance in overall performance in both the Massed condition, $b = 0.80$, 95% Wald CI = [0.43, 1.17], Wald $\chi^2(1) = 18.13$, $p < .001$, and the Interleaved condition, $b = 0.63$, 95% Wald CI = [0.27, 0.99], Wald $\chi^2(1) = 11.52$, $p < .001$, showing successful learning in both groups.

Comparison of massed, interleaved, and unstructured conditions

To further understand differences in performance among all three conditions, we conducted two additional analyses: one that assessed the effects of condition (with all three included) and a second one that compared Experiment 1 with Experiment 2. Note that these analyses were not preregistered. First, we fit a logistic mixed-effects model with a fixed effect of condition (Unstructured, Massed, or Interleaved; dummy coded) as well as by-item and by-participant random intercepts for condition. The results showed a marginal effect of condition, Wald $\chi^2(2) = 5.09$, $p = .08$. Pairwise comparisons using dummy coding revealed that participants were more accurate in the Massed condition than in the Unstructured condition, $b = 0.53$, Wald 95% CI = [0.06, 1.00], Wald $\chi^2(1) = 4.97$, $p = .03$. There was no significant difference between the Interleaved and Unstructured conditions, $b = 0.36$, Wald 95% CI = [−0.10, 0.83], Wald $\chi^2(1) = 2.34$, $p = .13$. In the second model, we compared Experiment 1 with Experiment 2 (collapsing across the Interleaved and Massed conditions). To do so, we fit a logistic mixed-effects model that included a fixed effect of experiment as well as by-item and by-participant random intercepts. The results showed a significant effect of experiment, with children in Experiment 2 performing significantly better than children in Experiment 1, $b = 0.45$, Wald 95% CI = [0.03, 0.86], Wald $\chi^2(1) = 4.51$, $p = .03$.

Discussion

Results from Experiment 2 revealed that children can successfully learn words given either massed or interleaved temporal structure. In addition, there was no significant difference between the Massed condition and the Interleaved condition. The lack of a condition effect in Experiment 2 and successful learning across both the Interleaved and Massed conditions point to children's flexibility in learning word–referent co-occurrences from massed or interleaved naming events.

However, we cannot strongly conclude that temporal structure does not affect children's learning. In particular, when examining all three conditions across Experiment 1 and Experiment 2, although there was no significant effect of condition, the mean for the Unstructured condition (.56) is notably lower than the means for the Interleaved (.64) and Massed (.67) conditions. Furthermore, although there was no difference between the Unstructured condition and the Interleaved condition, children did perform significantly worse in the Unstructured condition when compared with the Massed condition alone and with the Massed and Interleaved conditions combined. These results hint at an influence of temporal structure on children's learning. Given that children did not learn in the Unstructured condition and they performed worse in the Unstructured condition than in the Interleaved and Massed conditions combined, a lack of temporal structure may be particularly detrimental to children's learning. We return to these findings in the General Discussion.

The lack of observed differences between the Interleaved condition and the Massed condition is consistent with previous findings that children can track both massed and interleaved objects within the same task (Vlach & DeBrock, 2017). However, this finding is inconsistent with previous research in other domains showing effects of temporal structure on children's word learning and retention (Childers & Tomasello, 2002; Schwab & Lew-Williams, 2016; Vlach et al., 2012). Given these previous findings, it is surprising that children in our task did not benefit more from the Massed condition.

Furthermore, previous studies in adults have suggested that naming events that are close in time benefit from cross-situational word learning (Onnis, Edelmann, & Waterfall, 2011; Smith, Smith, & Blythe, 2011; Yurovsky & Frank, 2015; Kachergis et al., 2009). This suggests that there may be developmental differences in how learners benefit from massed naming events. In Experiment 3, we examined the possibility of developmental differences in how the temporal dynamics of naming events affect cross-situational word learning by testing a group of adults on the same materials that children were tested on in Experiments 1 and 2 (Unstructured, Massed, and Interleaved conditions).

Experiment 3

In Experiment 3, we aimed to determine how the structure of naming events affects adults' cross-situational word learning. To do so, we conducted a preregistered study (https://osf.io/84trw) in which we presented adults with the unstructured, massed, and interleaved tasks previously presented to children. The key question was whether adults learn the 8 word–object pairings differently across the three training structures and, furthermore, whether the training structure affects adults' word learning differently than it does that of children.

Method

Participants

A total of 90 undergraduates were recruited from the Arizona State University Introduction to Psychology participant pool. The sample included 35 women, all but 1 participant reported that they were native English speakers, and 14 participants indicated that they had knowledge of a second language in addition to English. Adults received course credit for participating, and informed consent was obtained for all participants according to the Arizona State University institutional review board.

Design

The design and materials were the same as in Experiments 1 and 2. Adults were randomly assigned to one of three conditions: Unstructured (n = 30), Massed (n = 30), or Interleaved (n = 30).

Procedure

Adults were asked to complete the task on a computer. The task was the same as that provided to children with a few minor changes. The experimenter's script was written out as instructions on the screen for participants to read. The auditory stimuli were presented over headphones, and adults were asked to use the computer mouse to indicate their responses. Furthermore, the familiarization phase was removed. After reading the instructions, adults were presented with training followed by testing. The entire task lasted approximately 5 min. After completing the task, adults filled out a language background questionnaire asking them details about their language history.

Results

The effect of trial structure on adults' test accuracy

Our first goal was to investigate whether our manipulation of temporal structure affected adults' accuracy at test. To test for an overall effect of condition, we conducted a preregistered analysis fitting a logistic mixed-effects model predicting adults' trial-by-trial accuracy from condition (dummy coded). We included only by-participant and by-item random intercepts after the model failed to converge with the addition of a by-item random slope for condition (this decision did not qualitatively affect the results). Including the fixed effect of condition significantly improved the fit of the model, Wald $\chi^2(2)$ = 23.32, p < .001, demonstrating a significant overall effect of condition on adults' test accuracy (see Fig. 3B).

We then used dummy coding to conduct pairwise comparisons of all levels of condition. Participants were more accurate in the Massed condition (M = .86, SD = .25, 95% CI = [.77, .95]) than both the Unstructured condition (M = .60, SD = .12, 95% CI = [.56, .65]), b = 1.99, Wald 95% CI = [1.17, 2.81], Wald $\chi^2(1)$ = 22.63, p < .001, and the Interleaved condition (M = .68, SD = .25, 95% CI = [.58, .77]), b = 1.50, Wald 95% CI = [0.69, 2.32], Wald $\chi^2(1)$ = 13.04, p < .001. Test accuracy did not significantly differ between the Interleaved and Unstructured conditions, b = 0.49, Wald 95% CI = [−0.22, 1.20], Wald $\chi^2(1)$ = 1.80, p = .18.

As a group, adults in the Unstructured condition did not show evidence of learning, b = 0.53, Wald 95% CI = [−0.22, 1.28], Wald $\chi^2(1)$ = 1.93, p = .16.[2] Adults were above chance for both the Massed con-

[2] Note that adults' accuracy in the Unstructured condition was above chance in analyses that do not consider by-item variability and instead simply summarize accuracy across all items (one-sample t test), $t(29)$ = 4.81, p < .001. The source of this discrepancy is that there is a high degree of by-item variability in the Unstructured condition, leading to a more conservative estimate of significance in the logistic mixed-effects models including by-item random effects. We discuss performance in the Unstructured condition in more detail in the General Discussion.

dition, b = 2.52, Wald 95% CI = [1.65, 3.40], Wald $\chi^2(1)$ = 31.87, p < .001, and the Interleaved condition, b = 1.02, Wald 95% CI = [0.25, 1.79], Wald $\chi^2(1)$ = 6.76, p = .009.

Comparing the effect of trial structure for adults versus children

Following our preregistered analysis plan, we next asked whether our manipulation of trial structure across condition had different effects on children's and adults' test accuracy. To test this question, we fit a logistic mixed-effects model predicting participants' trial-by-trial test accuracy from condition (Unstructured, Interleaved, or Massed; dummy coded), age group (adults or children), and their interaction. We included by-participant and by-item random intercepts after removing a by-item random slope for condition due to nonconvergence (including the slope does not qualitatively change the results). There was a significant interaction between condition and age group, Wald $\chi^2(2)$ = 10.61, p = .005, suggesting that the manipulation of trial structure affected children's and adults' test accuracy differently. The source of the interaction was that the difference between the Massed condition and the Unstructured condition, b = 1.13, Wald 95% CI = [0.37, 1.90], $\chi^2(1)$ = 8.47, p = .004, and the difference between the Massed condition and the Interleaved condition, b = 1.09, Wald 95% CI = [0.34, 1.84], $\chi^2(1)$ = 8.08, p = .004, were greater for adults than for children. Furthermore, adults' accuracy in the Massed condition (M = .86, SD = .25) was much higher than children's accuracy in the Massed condition (M = .67, SD = .22), b = 1.38, Wald 95% CI = [0.82, 1.95], Wald $\chi^2(1)$ = 22.80, p < .001. Interestingly, children and adults performed more similarly in the Interleaved condition (adults: M = .68, SD = .25; children: M = .64, SD = .21) and unstructured condition (adults: M = .60, SD = .12; children: M = .56, SD = .22), ps > .24.

Discussion

Results from Experiment 3 show that the structure of naming events affected adults' cross-situational word learning. Adults in the Massed condition performed the best at the task, outperforming adults in the Unstructured and Interleaved conditions, demonstrating that massed naming events provide an advantage for adults' word learning under ambiguity. When comparing adults' performance with that of children, analyses revealed a significant interaction between group and condition. Specifically, the performance advantage of adults in the Massed condition was not present in children. We discuss all results together in the General Discussion.

General discussion

In this set of experiments, we examined the role of temporal structure in cross-situational word learning. Children and adults experienced the same learning task; however, the temporal structure of naming events was manipulated such that trials—otherwise identical for all participants—were presented in an unstructured, interleaved, or massed order. Children in the Massed and Interleaved conditions were successful at learning, with no differences between these two conditions, but showed no evidence of learning in the Unstructured condition. By contrast, adults in the Massed condition performed the best when compared with adults in the other two conditions and when compared with children. Together, our findings demonstrate that adults significantly benefit from a massed temporal structure, whereas children do not. Furthermore, this effect was not simply due to adults being better at the task overall; children and adults performed similarly in the Unstructured and Interleaved conditions. The key difference in accuracy between the two groups was when naming events were presented in a massed structure.

Although the role of temporal structure on cross-situational word learning has been previously examined in a limited set of studies in children and adults (Onnis, Edelmann, & Waterfall, 2011; K. Smith, Smith, & Blythe, 2011; L.B. Smith & Yu, 2013; Vlach & Johnson, 2013; Yurovsky & Frank, 2015; Kachergis et al., 2009), our study compared children and adults directly, using the same task, across multiple between-participant temporal conditions. Thus, the current experiments are the first to examine the effects of temporal structure developmentally without confounds of task difficulty, structure type, and potential interactions caused by within-participant manipulations. Our experiments also focused on testing learners immediately after training, after only a few learning events (three exposures to each object), with a small set of words per trial (2 words and 2 objects on each trial), with a small set of total words (8 novel words overall), and manipulated the number of interleaving trials within a limited range (0–6 trials). Given that previous research has indicated that each of these factors influences cross-situational word learning and the effects of temporal structure (Cepeda et al., 2006; Vlach & Sandhofer, 2014; Yurovsky & Frank, 2015), it is possible that testing learners after a delay, increasing the number of exposures and items, and increasing our range of interleaving trials may yield different effects. For example, in adults, increasing the number of words presented on each trial, and therefore increasing the amount of ambiguity in a given learning instance, negatively affects cross-situational word learning performance (Yu & Smith, 2007). Does increasing the number of words presented on each trial strengthen or weaken the effects of massed structure on cross-situational word learning in adults and children? How the task demands interact with the effects of temporal structure during cross-situational word learning is an open question. Nonetheless, our findings are novel and clear; our sample of children and adults did not benefit in the same way from temporal structure. These findings highlight the need to better understand the different conditions under which temporal structure may affect children's and adults' word learning in future research.

The benefit for massed structure found in adults is consistent with previous research in learning and memory paradigms demonstrating a benefit of massed learning events for the immediate encoding of materials (Cepeda et al., 2006) as well as with research in other statistical learning domains showing that immediate repetitions boost adults' statistical learning (Onnis, Waterfall, & Edelman, 2008). Furthermore, our results confirm previous findings suggesting that adults' cross-situational word learning benefits from repetitions of words presented close in time (Onnis, Edelmann, & Waterfall, 2011; Smith, Smith, & Blythe, 2011; Yurovsky & Frank, 2015; Kachergis et al., 2009). Critically, because our study included multiple temporal conditions, it makes clear that massed naming events provide an advantage in adults for cross-situational word learning *over and above* interleaved or unstructured naming events.

In contrast, children were successful at learning from both massed and interleaved temporal structure. Our findings add to the limited research examining the effects of temporal structure on young children's cross-situational word learning over age. One previous study showed that children aged 2–5 years perform equally well with massed and interleaved objects presented within the same cross-situational word learning task (Vlach & DeBrock, 2017). On the other hand, younger children (12- to 16-month-olds) have been found to be affected by temporal structure when learning words under ambiguity (Smith & Yu, 2013; Vlach & Johnson, 2013). In particular, whereas 16-month-olds were successful at cross-situational word learning for objects presented with massed structure but not interleaved structure, 20-month-olds were successful at both types of structure (Vlach & Johnson, 2013). Therefore, our findings provide converging evidence that older children, who could be considered more sophisticated word learners, are able to flexibly learn from both massed and interleaved structure.

Our findings, however, differ from studies on ostensive word learning demonstrating a benefit for immediate encoding from massed object naming events (Childers & Tomasello, 2002; Schwab & Lew-Williams, 2016; Vlach et al., 2012). An open question is why children benefit from massed structure when learning from ostensive naming events but not when learning from ambiguous naming events. It is likely that the differing effects of temporal structure across these two learning situations have to do with the additional demands placed on learning when naming events are ambiguous. Future research can compare these two learning situations directly to examine whether indeed there are differential impacts of temporal structure on word learning with and without ambiguity.

The differences between children and adults in our task suggest that children do not exploit massed structure during cross-situational word learning to the same extent as adults do. These findings point to differences in the mechanisms underlying cross-situational word learning in children compared with adults. Although several researchers have put forward proposals describing the mechanisms driving cross-situational word learning, many of these proposals have been developed using adult data (Berens et al., 2018; Kachergis et al., 2012, 2017; Smith et al., 2011; Stevens et al., 2017; Trueswell et al., 2013; Yu & Smith, 2012; Yurovsky & Frank, 2015). Based on the current findings, however, we cannot assume that the way adults approach cross-situational word learning is the way children approach this task. Consistent with this notion, recent research has found that correct word–referent mappings on trials presented early in training are positively related to cross-situational word learning in 4-year-olds but are inversely related to learning in adults (Fitneva & Christiansen, 2017). In addition, 2.5-year-olds are more negatively affected than adults by distractors that co-occurred frequently with words during cross-situational word learning (Bunce & Scott, 2017). These studies, together with our results, suggest that adults and children indeed approach cross-situational word learning differently.

One source of developmental differences in tracking words across ambiguous situations may be found in attention and memory abilities in children compared with adults. A number of proposals agree that cross-situational word learning is subject to attention and memory constraints (Kachergis et al., 2012; Smith et al., 2011; Yu & Smith, 2012; Yurovsky & Frank, 2015). Time and repetition organize how memories are strengthened and how attention is allocated across learning events. For example, the memory of the association between a word and a referent is likely to weaken over time. An immediate repetition of the same word and object can strengthen an association before it has decayed in memory. Similarly, immediate repetitions may allow for comparison or priming processes to take effect, facilitating the connection of memories across learning events and also highlighting the familiar from the novel (Kachergis et al., 2009). All of these factors have been proposed as possible mechanisms by which the distribution of learning events influences memories (Cepeda et al., 2006; Goldstein et al., 2010). The impact of memory and attentional processes at different developmental timepoints may also interact with the complexity of the learning problem that children and adults face (Yurovsky & Frank, 2015). For example, learners' ability to derive benefits from massed structure may depend on how easily they can maintain multiple options in memory. For adults, the relatively small set of novel words (8 total words) in the current experiment may have placed limited demands on working memory, allowing them to explicitly track candidate words across successive training trials. On the other hand, children may have had more difficulty in tracking multiple object–label connections explicitly for the current set size, hindering them from exploiting the massed structure to the same degree as adults. Understanding how attention and memory processes affect cross-situational word learning—at different levels of complexity—and how these effects change across development is critical to understanding how word learning unfolds over time.

As a first step in this direction, a recent study set out to examine how memory abilities predict cross-situational word learning in preschoolers (Vlach & DeBrock, 2017). The results showed not only that memory abilities accounted for cross-situational word learning abilities over and above age and vocabulary abilities but also that different kinds of memory abilities (children's individual abilities to remember words, objects, and word–object bindings) positively predicted cross-situational word learning. What these findings suggest is that as children's memory abilities improve, so does their cross-situational word learning. However, if memory abilities are limited in children, then the question remains why children did not benefit *more* from massed structure than adults given that massed structure is often found to reduce memory demands. One possibility, as suggested above, is that children's ability to explicitly attend to specific object–label pairings across successive trials was limited in the current task, perhaps due to the difficulty of tracking, at least from children's perspective, a relatively large set of novel items. Another possibility is that the temporal structure of learning events interacts with the timing of when learners are tested (Cepeda et al., 2006). Past work suggests that although massed training structure may at times be beneficial over short retention intervals, the benefit of more widely spacing learning events increases with longer retention intervals (Carvalho & Goldstone, 2014, 2017; Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008). In the current work, word learning was always tested immediately following the training procedure, that is, at a short retention interval. However, children may have experienced more forgetting during the delay between training and test compared with adults. In this case, the interleaved training may have led to similar performance compared with the massed condition for children because it helped to counteract the rapid decay in memory for novel object names, whereas adults experienced little or no decline in memory during the short retention interval. Thus, based on the current findings, a fruitful avenue for future research is to understand how improvements in attention and memory abilities over early childhood may lead to the successful exploitation of massed structure for cross-situational word learning in short-term retention and what benefits interleaved training may bring for long-term retention.

Although there was a notable difference in how children and adults exploited massed structure to learn novel words in our task, a similarity between child and adult performance was that neither group successfully learned in the Unstructured condition. Although the Unstructured condition was the same as the Interleaved condition in the average number of trials in between repetitions of an object, it also had a wider range of trials between repetitions of objects (1–6) than the other two conditions (Massed: 0–2; Interleaved: 2–4). Having so many interleaving trials between repetitions of some objects (up to 6) may have made the task much more difficult for learners. It is possible that a larger number of trials presented in between repetitions of objects hinders performance (Yurovsky & Frank, 2015).

There are, however, several additional reasons for the lack of significant learning in the Unstructured condition. Children's accuracy scores for the Unstructured condition are comparable to means found in previous research in this age group showing evidence of successful learning (see Fitneva & Christiansen, 2017, and Vlach & DeBrock, 2017), suggesting that Experiment 1 might not have had sufficient power to detect learning at the group level for children. Adults' accuracy is lower than previously found under this low-ambiguity context (see Yu & Smith, 2007), but we note that our task did present fewer training trials. Given that learners' accuracy does improve with more word–referent repetitions (e.g., Zettersten, Wojcik, Benitez, & Saffran, 2018), our task may have been more difficult given the low number of training trials per word. Finally, there was a great deal of by-item variability in adult performance, suggesting that the particular unstructured orders created to test learners in our task may have been difficult.

Nonetheless, adults' and children's poor performance in our Unstructured condition raises the question of how structure, as compared with a random presentation of naming events, may benefit ambiguous word learning over time. In particular, naming events in the real world often contain structure, not just temporally (Smith et al., 2018) but also with respect to the spatial and linguistic contexts in which objects and words appear (Roy, Frank, DeCamp, Miller, & Roy, 2015). Furthermore, caregivers structure their naming events to include repetition and predictable regularities that direct attention to the right referent at the right time during naming (Chang & Deák, 2019; Schwab et al., 2018; Yu & Smith, 2017). For young children, predictable events, consistent actions, consistent locations, and consistent syntactic frames benefit word learning (Benitez & Saffran, 2018; Benitez & Smith, 2012; Eiteljoerge, Adam, Elsner, & Mani, 2019; Vlach & Sandhofer, 2011; Wojcik & Saffran, 2015). These different kinds of structure are also likely to boost cross-situational word learning. Therefore, it is notable that most studies in word learning take pains to remove much of this structure in their designs, for example, carefully randomizing the temporal order of training trials. Manipulating different types of structure, including the distribution of naming events, and conducting naturalistic studies to determine the kinds of structure present across real-world naming events will help to better elucidate how children exploit different types of regularities during word learning over time.

In conclusion, we set out to examine how temporal structure affects children's and adults' cross-situational word learning. We tested both groups of learners under the same conditions and manipulated the temporal structure of naming events to be unstructured, interleaved, or massed. Results showed that whereas children and adults performed similarly for unstructured and interleaved naming events, adults benefited more dramatically from massed structure compared with children. These results demonstrate that temporal structure differentially affects children's and adults' cross-situational word learning and highlights the need to better understand how children and adults track word–referent statistics over time.

Acknowledgments

This work was supported in part by a National Science Foundation grant (NSF-GRFP DGE-1747503) awarded to M.Z. We thank Marianna Kaneris, Emily Fatula, and Cassandra Leedom for helping to recruit and test our child participants as well as Harrison Westerfield and Elaina Corbin for helping to recruit and test adults. Special thanks go to the staff members at the Children's Museum of Phoenix for their support during data collection and to all the parents and children who participated. All data,

analysis scripts, and experiment materials are available in public repositories on GitHub (https://github.com/mzettersten/crossSitRep) and Open Science Framework (https://osf.io/2hmxr/).

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jecp.2020.104961.

References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.
Bates, D., Mächler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48.
Benitez, V. L., & Saffran, J. R. (2018). Predictable events enhance word learning in toddlers. *Current Biology, 28*, 2787–2793.
Benitez, V. L., & Smith, L. B. (2012). Predictable locations aid early object name learning. *Cognition, 125*, 339–352.
Berens, S. C., Horst, J. S., & Bird, C. M. (2018). Cross-situational learning is supported by propose-but-verify hypothesis testing. *Current Biology, 28*, 1132–1136.
Blythe, R. A., Smith, K., & Smith, A. D. (2010). Learning times for large lexicons through cross-situational learning. *Cognitive Science, 34*, 620–642.
Bunce, J. P., & Scott, R. M. (2017). Finding meaning in a noisy world: Exploring the effects of referential ambiguity and competition on 2.5-year-olds' cross-situational word learning. *Journal of Child Language, 44*, 650–676.
Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development, 15*, 17–29.
Carvalho, P. F., & Goldstone, R. L. (2014). Effects of interleaved and blocked study on delayed test of category learning generalization. *Frontiers in Psychology, 5*. https://doi.org/10.3389/fpsyg.2014.00936.
Carvalho, P. F., & Goldstone, R. L. (2017). The sequence of study changes what information Is attended to, encoded, and remembered during category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43*, 1699–1719.
Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin., 132*, 354–380.
Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science, 19*, 1095–1102.
Chang, L. M., & Deák, G. O. (2019). Maternal discourse continuity and infants' actions organize 12-month-olds' language exposure during object play. *Developmental Science, 22* e12770.
Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology, 38*, 967–978.
Eiteljoerge, S. F. V., Adam, M., Elsner, B., & Mani, N. (2019). Word–object and action–object association learning across early development. *PLoS ONE, 14*(8) e220317.
Fitneva, S. A., & Christiansen, M. H. (2017). Developmental changes in cross-situational word learning: The inverse effect of initial accuracy. *Cognitive Science, 41*, 141–161.
Frank, M. C., Tenenbaum, J. B., & Fernald, A. (2013). Social and discourse contributions to the determination of reference in cross-situational word learning. *Language Learning and Development, 9*, 1–24.
Goldstein, M. H., Waterfall, H. R., Lotem, A., Halpern, J. Y., Schwade, J. A., Onnis, L., & Edelman, S. (2010). General cognitive principles for learning structure in time and space. *Trends in Cognitive Sciences, 14*, 249–258.
Halberda, J. (2003). The development of a word-learning strategy. *Cognition, 87*, B23–B34.
Horst, J. S. (2013). Context and repetition in word learning. *Frontiers in Psychology, 4*. https://doi.org/10.3389/fpsyg.2013.00149.
Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) database: A collection of novel images for use in experimental research. *Behavior Research Methods, 48*, 1393–1409.
Ichinco, D., Frank, M., & Saxe, R. (2009). Cross-situational word learning respects mutual exclusivity. In N. A. Taatgen & H. van Rijn (Eds.), Proceedings of the 31st annual meeting of the Cognitive Science Society (pp. 2214–2219). Austin, TX: Cognitive Science Society.
Kachergis, G., Shiffrin, R., & Yu, C. (2009). Temporal contiguity in cross-situational statistical learning. In N. A. Taatgen & H. van Rijn (Eds.), Proceedings of the 31st annual meeting of the Cognitive Science Society (pp. 1704–1709). Austin, TX: Cognitive Science Society.
Kachergis, G., Yu, C., & Shiffrin, R. M. (2012). An associative model of adaptive inference for learning word–referent mappings. *Psychonomic Bulletin & Review, 19*, 317–324.
Kachergis, G., Yu, C., & Shiffrin, R. M. (2017). A bootstrapping model of frequency and context effects in word learning. *Cognitive Science, 41*, 590–622.
Kucker, S. C., McMurray, B., & Samuelson, L. K. (2015). Slowing down fast mapping: Redefining the dynamics of word learning. *Child Development Perspectives, 9*, 74–78.
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 of the United States of America, 108*, 9014–9019.
Onnis, L., Edelmann, S., & Waterfall, H. (2011). Local statistical learning under cross-situational uncertainty. InL. Carlson (Ed.), Proceedings of the 33rd annual meeting of the Cognitive Science Society. Austin, TX: Cognitive Science Society, pp. 2697–2702..
Onnis, L., Waterfall, H. R., & Edelman, S. (2008). Learn locally, act globally: Learning language from variation set cues. *Cognition, 109*, 423–430.
Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.

R Development Core Team. (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.r-project.org/.
Regier, T. (2005). The emergence of words: Attentional learning in form and meaning. *Cognitive Science, 29*, 819–865.
Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences of the United States of America, 112*, 12663–12668.
Schwab, J. F., & Lew-Williams, C. (2016). Repetition across successive sentences facilitates young children's word learning. *Developmental Psychology, 52*, 879–886.
Schwab, J. F., Rowe, M. L., Cabrera, N., & Lew-Williams, C. (2018). Fathers' repetition of words is coupled with children's vocabularies. *Journal of Experimental Child Psychology, 166*, 437–450.
Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition, 61*, 39–91.
Smith, K., Smith, A. D., & Blythe, R. A. (2011). Cross-situational learning: An experimental study of word-learning mechanisms. *Cognitive Science, 35*, 480–498.
Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences, 22*, 325–336.
Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word–referent learning. *Trends in Cognitive Sciences, 18*, 251–258.
Smith, L., & Yu, C. (2008). Infants rapidly learn word–referent mappings via cross-situational statistics. *Cognition, 106*, 1558–1568.
Smith, L. B., & Yu, C. (2013). Visual attention is not enough: Individual differences in statistical word–referent learning in infants. *Language Learning and Development, 9*, 25–49.
Stevens, J. S., Gleitman, L. R., Trueswell, J. C., & Yang, C. (2017). The pursuit of word meanings. *Cognitive Science, 41*, 638–676.
Suanda, S. H., Mugwanya, N., & Namy, L. L. (2014). Cross-situational statistical word learning in young children. *Journal of Experimental Child Psychology, 126*, 395–411.
Suanda, S. H., Smith, L. B., & Yu, C. (2016). The multisensory nature of verbal discourse in parent–toddler interactions. *Developmental Neuropsychology, 41*, 324–341.
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*, 126–156.
Vlach, H. A., Ankowski, A. A., & Sandhofer, C. M. (2012). At the same time or apart in time? The role of presentation timing and retrieval dynamics in generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 246–254.
Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children's cross-situational word learning, memory, and language abilities. *Journal of Memory and Language, 93*, 217–230.
Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition, 127*, 375–382.
Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology, 108*, 394–401.
Vlach, H. A., & Sandhofer, C. M. (2014). Retrieval dynamics and retention in cross-situational word learning. *Cognitive Science, 38*, 757–774.
Wojcik, E. H., & Saffran, J. R. (2015). Toddlers encode similarities among novel words from meaningful sentences. *Cognition, 138*, 10–20.
Yu, C., & Ballard, D. H. (2007). A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing, 70*, 2149–2165.
Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science, 18*, 414–420.
Yu, C., & Smith, L. B. (2012). Modeling cross-situational word–referent learning: Prior questions. *Psychological Review, 119*, 21–39.
Yu, C., & Smith, L. B. (2017). Hand–eye coordination predicts joint attention. *Child Development, 88*, 2060–2078.
Yurovsky, D., & Frank, M. C. (2015). An integrative account of constraints on cross-situational learning. *Cognition, 145*, 53–62.
Zettersten, M., Wojcik, E., Benitez, V. L., & Saffran, J. (2018). The company objects keep: Linking referents together during cross-situational word learning. *Journal of Memory and Language, 99*, 62–73.