Published in final edited form as:
Cogn Sci. 2021 July ; 45(7): e13010. doi:10.1111/cogs.13010.

Cross-situational Learning From Ambiguous Egocentric Input Is a Continuous Process: Evidence Using the Human Simulation Paradigm

Yayun Zhang[a], **Daniel Yurovsky**[b], **Chen Yu**[a,c]

[a]Department of Psychology, University of Texas, Austin

[b]Department of Psychology, Carnegie Mellon University

[c]Department of Psychological and Brain Sciences, Indiana University, Bloomington

Abstract

Recent laboratory experiments have shown that both infant and adult learners can acquire word-referent mappings using cross-situational statistics. The vast majority of the work on this topic has used unfamiliar objects presented on neutral backgrounds as the visual contexts for word learning. However, these laboratory contexts are much different than the real-world contexts in which learning occurs. Thus, the feasibility of generalizing cross-situational learning beyond the laboratory is in question. Adapting the Human Simulation Paradigm, we conducted a series of experiments examining cross-situational learning from children's egocentric videos captured during naturalistic play. Focusing on individually ambiguous naming moments that naturally occur during toy play, we asked how statistical learning unfolds in real time through accumulating cross-situational statistics in naturalistic contexts. We found that even when learning situations were individually ambiguous, learners' performance gradually improved over time. This improvement was driven in part by learners' use of partial knowledge acquired from previous learning situations, even when they had not yet discovered correct word-object mappings. These results suggest that word learning is a continuous process by means of real-time information integration.

Keywords

Word learning; Early language acquisition; Statistical learning; Cross-situational learning

1. Introduction

Children learn words in a complex environment. When hearing a word, children need to infer the meaning of the word despite the uncertainty about its potential referent (Quine, 1960). Many heuristics have been shown to help reduce referential uncertainty at the moment, such as attentional cues provided by the speaker (Baldwin, 1991, 1993), the whole-object assumption (Macnamara, 1972), and mutual exclusivity (Halberda, 2003; Golinkoff et al, 1992; Markman & Wachtel, 1988). However, the degree of referential uncertainty

Correspondence should be sent to Yayun Zhang, Department of Psychology, University of Texas at Austin, 108 E Dean Keeton St, Austin, TX 78712. yayunzhang@utexas.edu.

can vary widely in everyday learning situations. While some situations are transparent, allowing young learners to easily identify the intended referent of a parent's speech (Bloom, 2000; Carey & Bartlett, 1978; Markman, 1989; Regier, 2005), others are highly ambiguous, requiring young learners to make a correct inference from multiple referent candidates. Recent studies show that young children are able to build a correct word-referent mapping not only by inferring a word's referent correctly in a single situation but also by aggregating statistical information across multiple individually ambiguous situations (Akhtar & Montague, 1999; Scott & Fisher, 2011; Smith & Yu, 2008, 2013; Vlach & Johnson, 2013; Vouloumanos & Werker, 2009).

The computational mechanism involved in this process of inferring a correct word-referent mapping from multiple learning situations has been called cross-situational learning (Siskind, 1996; Smith et al., 2011; Yu & Smith, 2007). The essential idea of cross-situational learning is this: because a label and its correct referent are likely to co-occur more consistently than incorrect pairs that happen to co-occur by chance, the correct mappings can be identified by tallying co-occurring statistics across multiple learning situations (Smith & Yu, 2008; Yu & Smith, 2007). Many experimental and modeling studies of cross-situational learning have shown that infants, children, and adults are able to use cross-situational statistics to acquire word-referent mappings (e.g., Blythe et al., 2010; Chen et al., 2018; Fitneva & Christiansen, 2011; Kachergis et al., 2012; Koehne & Crocker, 2015; Monaghan et al., 2015; Onnis et al., 2011; Rebuschat et al., 2021; Smith et al., 2011; Smith & Yu, 2008; Suanda & Namy, 2012; Trueswell et al., 2013; Wang & Mintz, 2018; Yurovsky et al., 2013; Yu & Smith, 2007). However, stimuli used in most experimental tasks are highly controlled—usually composed of a sequence of trials, each of which has a small number of unfamiliar objects separated from each other in space on a monochromatic background. Compared with the experimental paradigms used in cross-situational learning studies, the visual environment that young children experience in the real world may be much more ambiguous, containing many referents onto which a word could possibly map. Therefore, a critical question that has been recently raised is whether the cross-situational learning solution demonstrated in well-controlled laboratory tasks can be generalized to word learning in the real world (Medina et al., 2011). The goal of the current study is to examine the feasibility of building word-referent mappings from a sequence of ambiguous situations extracted from real-world contexts and investigate the learning mechanisms that operate over them. Toward these goals, we used the Human Simulation Paradigm (HSP) originally introduced in Gillette et al. (1999) as an experimental paradigm to incorporate naturalistic learning contexts in the real world to cross-situational learning experimental tasks in the laboratories.

In the HSP, naming instances from naturalistic parent–child interactions are segmented and extracted as individual vignettes. The audio of each vignette is muted, and a beep is inserted at the point in time corresponding to the parent's production of an object's label. Learners are then presented with a sequence of vignettes and asked to guess which referent's label was produced in each vignette. Medina et al. (2011) used the HSP to investigate cross-situational learning by testing whether adults are able to accumulate statistical evidence from a sequence of vignettes extracted from parent–child interactions. They found no evidence of incremental learning from multiple ambiguous learning situations. Instead, successful word learning in their study depended solely on the presence of unambiguous learning situations. Moreover, participants learned the best when the unambiguous learning situations happened early in training, suggesting that word learning requires an initial "one-shot learning" step followed by confirmation (Trueswell et al., 2013). This Propose-but-Verify (PbV) learning process has been further supported by a subsequent experimental and computational modeling studies (i.e., Stevens et al., 2017). Together, those studies suggest that cross-situational learners may form an initial hypothesis about the meaning of a word on their first encounter with it, and then rely on subsequent encounters to either confirm or reject hypothesized mappings.

Yurovsky et al. (2013), however, found contrasting results using the same HSP method. In their study, vignettes were extracted from videos captured by a camera mounted on children's heads. They used these egocentric vignettes to study the effect of the child's first-person perspective on the process of cross-situational learning. Contrary to Medina et al.'s findings, participants' learning performance improved significantly after watching multiple vignettes from the child's view but not after viewing the same naming instances recorded from a third-person view (Yurovsky et al., 2013). These results suggest that the information available in first-person views may recruit integrative processes not involved in learning from third-person views.

The studies reported in Medina et al. (2011) and Yurovsky et al. (2013) used the same experimental paradigm; and they both found evidence of successful learning. However, the conclusions from the two studies differ, raising several critical questions about cross-situational learning. First, can learners acquire correct word-referent mappings when all learning situations are ambiguous? Both studies presented learners with a mix of ambiguous and unambiguous naming instances, and learners may have relied entirely on the unambiguous naming instances in the mix. In the present study, we exclude unambiguous naming instances and focus on learning solely from ambiguous situations. We predict that even when individual learning instances are ambiguous in isolation, learners will aggregate information across them to converge to correct word-referent mappings.

Second, how do statistical learners use prior knowledge acquired from previous situations to resolve ambiguity in a new learning situation? The correct information learners gathered in the past would certainly help learning by reducing the degree of uncertainty in subsequent situations. However, given that learners cannot always obtain the right information from ambiguous situations, an open question is whether imperfect statistical information gathered from previous situations would still contribute (or hurt) subsequent learning. In other words, is partial knowledge still helpful in cross-situational learning? We predict that partial information from ambiguous learning situations would improve statistical learning if learners were able to continuously integrate statistical information from ambiguous situations.

The third question concerns the amount of data used in cross-situational learning. Some have argued that the key computational mechanism of cross-situational learning is built on hypothesis testing in which learners propose and confirm specific hypotheses (Medina et al., 2011; Trueswell et al., 2013; Stevens et al., 2017). This type of statistical computation relies heavily on integrating information from present and immediate past experiences to confirm or reject newly formed hypotheses. Another account, based on associative learning, views statistical learning not as all-or-none in hypothesis testing, but as a continuous process through which statistical evidence gradually accumulated across multiple situations (Kachergis et al., 2012; Yu & Smith, 2007). Under this account, all past experience, either immediate or distant, matters to learning. In the present study, we conducted several detailed analyses to examine whether cross-situational learning is built upon all the past learning situations or just the immediate past.

To answer these questions, we designed three experiments using the HSP. Experiment 1 was designed to provide a baseline for the following two cross-situational learning experiments by quantifying the ambiguity of individual naming instances. In this experiment, participants were provided with individual naming instances and asked to guess a target referent for each instance. Experiment 2 used the highly ambiguous naming events identified in Experiment 1 to test cross-situational learning exclusively from ambiguous learning instances. Participants were exposed to a sequence of learning trials referring to the same target referent and asked to guess the target referent trial by trial. Using participants' trial-by-trial responses, we examined how participants' responses in previous trials influence their responses in the current trial. One potential limitation of Experiment 2 is that the process of accumulating statistical information was interwoven with retrieval requests—participants were asked to guess after each trial. If this retrieval process affects how statistical information is processed or stored, then the learning mechanisms discovered in Experiment 2 may not generalize to more naturalistic situations in which learners are not constantly asked to retrieve what they know (Karpicke & Roediger, 2008; Roediger & Butler, 2011; Vlach & Johnson, 2013). Experiment 3 was designed to address this issue by using a similar training session as that used in Experiment 2, except that learning accuracy was measured only at the end of the training session. Experiment 3 allowed us to directly compare the overall learning accuracy in a continuous learning mode, to that from Experiment 2 in which participants were asked to retrieve their current knowledge trial by trial. Similar results from the two experiments would suggest that the findings from trial-by-trial data in Experiment 2 can also be applied to explain cross-situational learning in a continuous and uninterrupted mode. Taken together, the three experiments aimed at providing new evidence on cross-situational learning in highly ambiguous situations, and importantly, shedding light on real-time mechanisms through which statistical information accumulates in cross-situational learning.

2. Experiment 1

The goal of Experiment 1 was to quantify the degrees of ambiguity of a set of naming instances occurring in free-flowing parent–child toy play. Based on this assessment, a subset of highly ambiguous learning situations from those vignettes were selected for Experiments 2 and 3. In total, 96 naming vignettes collected by Yurovsky et al. (2013) were presented to a group of adult learners who were asked to guess the target referent of each vignette as a way to assess the degree of uncertainty of those naming instances.

2.1. Method

2.1.1. Participants—Seventeen Indiana University undergraduates (M_{age} = 19.82 years, SD_{age} = 1.47 years) participated in exchange for course credits. All participants completed the experiment.

2.1.2. Materials—The video corpus in Yurovsky et al. (2013) included eight play sessions, in which each parent–child dyad was asked to play with 25 toys as they naturally would at home for 10 minutes. The play interaction was recorded from two views: a third-person view from a tripod camera and a first-person view from a camera mounted on the child's head. We used the videos recorded from the child's first-person view in the present study because the visual information from this egocentric view is a close approximation of the relevant information that the child learner perceives and uses in statistical learning (Yu et al., 2009).

Ninety-six vignettes of naming instances were extracted from the video corpus. The target referents were 12 unique toys (e.g., elephant, mickey, tiger, etc.), each of which had eight naming instances from at least four different parent–child dyads. Those naming vignettes appeared to vary in their degrees of ambiguity. Some instances (e.g., Fig. 1a) are unambiguous with a single dominant object in the child's view and others (e.g., Fig. 1b) are highly ambiguous with several potential referents in view. We grouped these 96 vignettes into eight blocks with 12 vignettes in each block referring to each of the 12 toys. Moreover, vignettes within each block were randomized. As a result, there were 12 (ranging between 5 and 19) trials between two vignettes referring to the same target. This design made it unlikely for learners to use previous responses to inform responses in subsequent trials, and therefore minimized the possibility that learners performed cross-situational learning in this baseline condition.

For each naming instance, the original sound was muted, and the toy's name was replaced by a beep at the onset of the label. Most vignettes were 5 seconds long, with the name's onset occurring at exactly the third second. Two more seconds were added to the vignettes if parents said the toy name again within 2 seconds after the first naming instance. Seven of the 96 vignettes included two naming instances and two included three naming instances. Four additional vignettes were used as training examples before the experiment to make sure participants understood the task. None of the correct referents in these examples were one of 12 targets. In addition to video vignettes, a forced-choice test was designed, which contained 25 color photographs of all candidate toys used in the parent–child free-play session. The photos were displayed in a 5 × 5 grid on a white background,

2.1.3. Procedure—Participants were instructed to watch the vignettes and guess which object was likely to be labeled by the parents. After seeing each vignette, they were asked to guess the most likely referent from 25 pictures by clicking on the guessed picture. No feedback was given. At the beginning of the study, participants were familiarized with the task through four "warm-up" vignettes, each followed by a testing trial. Once they were familiar with the study procedure, they were prompted to begin the actual experiment.

2.2. Results and discussion

As shown in Fig. 2, individual naming instances extracted from free-flow toy play showed a wide range of ambiguity. Some naming instances seemed to be highly unambiguous as all the participants responded correctly; other instances were highly ambiguous as none of the participants responded correctly; and the rest contained various degrees of ambiguity as only some participants were able to choose the correct target from the 25 candidate toys provided at test. This distribution seems to reflect the variability of parent naming that is expected to observe from natural parent–child interaction. When parents named toy objects in free play, they sometimes named them at the exact moments when the target object was dominant in the child's view and therefore was likely to be considered as the referent of a name. But in other moments, spontaneous naming from parents may be out of synch with the child's attention as the child might attend to more than one object or even worse, to a single object that was not the one named by the parent (Zhang & Yu, 2017). As shown in Fig. 2, roughly 40% of vignettes were unambiguous, leading participants to guess accurately more than 70% of the time ($M = 0.94$, $SD = 0.07$, $Min = 0.71$; $Max = 1$). For the remaining 60% of vignettes, participants' response accuracies were well below 70% ($M = 0.14$, $SD = 0.16$, $Min = 0$; $Max = 0.59$). We classified these as ambiguous. Based on this criterion, 60 out of the 96 ambiguous naming instances tested were used in the following experiments to examine whether learners could aggregate statistical evidence from those ambiguous trials. Note that these ambiguous vignettes still varied in their ambiguity, from naming events in which no participants guessed correctly, to those in which a large fraction of participants did guess the correct answer.

In summary, naming events during naturalistic parent–child joint play vary in ambiguity. A subset of highly ambiguous naming instances was selected and used in the following experiments to examine whether learners can accumulate cross-situational statistics from those ambiguous situations. Toward this goal, accuracy measures on individual naming instances from Experiment 1 were also used as baselines in the following experiments.

3. Experiment 2

Experiment 2 was designed to examine cross-situational learning from ambiguous learning events. Specifically, we aimed at answering two questions: can learners find the correct referent after being exposed to a sequence of ambiguous learning situations? If so, what mechanisms do they use to aggregate statistical evidence trial by trial in real-time learning?

3.1. Method

3.1.1. Participants—Twenty-six Indiana University undergraduates ($M_{\text{age}} = 19.08$ years, $SD_{\text{age}} = 1.20$ years) participated and received course credits. None had participated in the previous baseline study or other cross-situational word learning experiments.

3.1.2. Materials—The sixty ambiguous trials selected from Experiment 1 were grouped into 12 blocks. Each block had five different ambiguous vignettes all referring to the same target. Each of the 12 target toys was assigned a novel two-syllable label (e.g., agen, gree, hage, etc.). These labels were recorded by a female native speaker of English. Instead of beeps, the novel labels were now inserted at the exact moment when parents named an object in a vignette. Within the five vignettes in a block, the same label was played as a clear indicator that the five naming instances shared the same target label. We used the same forced-choice test procedure used in Experiment 1.

3.1.3. Procedure—The procedure was similar to the one used in Experiment 1. The key difference was that participants were told that all five vignettes within a block referred to the same target. This instruction was further enforced by providing the same label within a block and different ones across blocks. Throughout the learning trials within a block, participants could change their response on any given trial. However, if they believed their previous answer was correct, they should stay with the same answer. They were not allowed to go back and change their previous answers and no feedback was provided through the whole experiment. Participants saw 12 blocks of trials in total. After each block, a prompt would appear to remind participants to get ready for the next block of trials.

3.2. Results and discussion

To measure whether participants accumulate knowledge across ambiguous naming instances, we calculated the response accuracies trial by trial. Fig. 3 shows both accuracy in the cross-situational learning condition in Experiment 2, and corresponding baseline accuracy of the same individual trials in Experiment 1.[1] Trial-by-trial learning performance clearly shows a dramatic improvement, from 23% accuracy on Trial 1 to almost 50% on Trial 5. To formally test the improvement over trials, we fit a mixed-effects logistic regression predicting accuracy from trial number and baseline accuracy from Experiment 1 while also taking into account the random intercepts for each subject (mixed effects model: accuracy ~ trial + baseline + (1|subject)). This model revealed a significant main effect of trial number ($\beta = 0.29$, $p < .001$) over and above the effect of baseline accuracy ($\beta = 2.42$, $p < .001$). This improvement observed in the present study supports the hypothesis that highly ambiguous instances alone are sufficient to produce successful learning. Even though the learning instances were individually ambiguous (on average, 14% accuracy for all 60 trials), they jointly created much less ambiguous data for learners who aggregated across them. The gradual trial-by-trial improvement suggests that word learning is a continuous process, wherein statistical learners make progress by integrating what they have learned from previous situations with the information presented in the current situation. For instance, compared with the 16% baseline, the 50% accuracy in Trial 5 results from integrating the information acquired in the first four situations with the information presented in Trial 5. We next report a set of detailed analyses on a trial-by-trial basis to reveal underlying statistical computations that contribute to the increasing improvement.

[1] Accuracy on the first trials ($M_1 = 0.23$, $SD_1 = 0.16$, 95% CI [0.18, 0.28]) were expected to be similar to baseline but still significantly higher ($M_{\text{baseline}} = 0.11$, $SD_{\text{baseline}} = 0.32$, 95% CI [0.07, 0.15], $\beta = 0.85$, $p < .01$). This is because the mean first-trial accuracy was calculated by aggregating responses across blocks and participants tended to achieve better learning performance in the first trials of later blocks, suggesting that learners may also learn cross-situationally across multiple target words. For example, if learners chose object A as the correct target in the first block, they could be less likely to choose object A again in later blocks as they were told to choose different toys in different blocks. In this way, learners adopted the mutual exclusivity strategy to narrow down their search space and improve response accuracy for later trials. Therefore, the first responses in later blocks had higher accuracy than the ones from earlier blocks because learners not only aggregated information within blocks but also accumulated statistics continuously throughout the entire study. The topic of cross-block statistical integration is worth future studies by itself. Nonetheless, the present study focuses on information aggregation from multiple learning instances within a block.

3.2.1. Effect of previously acquired knowledge on subsequent learning—One fundamental mechanism in statistical learning is to remember and use what has been learned in the past to improve learning in the future (e.g., Thiessen, 2017; Yurovsky et al., 2014). In the present study, we considered two scenarios defined by whether previously acquired knowledge is correct or incorrect. We first calculated participants' learning performance on the current trial conditioned on whether their previous response was correct. As shown in Fig. 4, when participants made a correct response from a previous trial, they were much more likely to stay with the correct answer ($M = 0.84$, $SD = 0.36$, 95% CI [0.81, 0.88]), compared to when they made an incorrect response from a previous trial ($M = 0.17$, $SD = 0.38$, 95% CI [0.18, 0.20]). To determine whether this difference was statistically significant, we fit a mixed effect model as before, but this time added an additional main effect of previous trial, which was coded as −1 if incorrect, and 1 if correct (mixed effects model: accuracy ~ trial + baseline + previous trial + (1|subject)). We found that all factors included in the previous model remained significant in the current model (trial number: $\beta = 0.31$, $p < .001$, baseline accuracy: $\beta = 2.85$, $p < .001$), and moreover the new factor of previous accuracy was also a significant predictor ($\beta = 1.48$, $p < .001$). If participants found the correct target in a previous trial, they were much likely to identify the same target again in the current trial to confirm their previous selection, instead of starting from scratch with an uninformed response. In this way, correct information previously acquired is integrated with the current information to improve subsequent learning.

However, with highly ambiguous situations used in the present study, the learners were more likely to make incorrect than correct responses at the beginning of the learning. If their previous responses were incorrect, their hypothesized meaning would be disconfirmed on the current trial. In those situations, would subsequent learning still benefit from incorrect responses made in the previous trials? To answer this question, we examined the trials preceded by incorrect responses and compared accuracies on these trials with the baseline accuracies obtained in Experiment 1 (mixed effects model: accuracy ~ baseline + (1|video) + (1|subject)). This model found a significant effect of experiment, indicating that even when participants failed to obtain the correct answer from a previous trial, they still had a better chance to be correct ($M = 0.17$, $SD = 0.38$, 95% CI [0.18, 0.20]) than they would have been with only the current trial's information ($M = 0.14$, $SD = 0.16$, 95% CI [0.13, 0.15], $\beta = 2.99$, $p < .01$). There are two possible learning mechanisms through which incorrect information from previous experiences can still help subsequent learning. First, after learners disconfirm an incorrect response in a prior trial, they could exclude the same wrong answer from consideration on the current trial (Yurovsky et al., 2014). Second, learners could potentially encode more information than just their single best guess on each trial (Frank et al., 2009; Smith & Yu, 2008; Vouloumanos, 2008; Yu et al., 2005; Yu & Smith, 2007). They could also store other word-object associations built from the co-occurrence statistics of previous trials. After discarding a wrong response, statistical learners could access stored associations and give a new response based not only on the current information but also the previously stored associations. This new response would have a better chance to be correct than an uninformed response. In this way, partial knowledge about potential word-referent mappings could guide the learners to make a more informed response in the current trial. This explanation contradicts the findings reported in Medina et al. (2011), showing that when participant responds incorrectly from a learning situation, their response accuracy is at chance at the very next learning situation, indicating no knowledge of previous contexts.

3.2.2. Effect of amount of information on subsequent learning—

For any statistical learning model, the amount of training data is a critical contributor to successful learning. However, two prominent accounts of cross-situational learning differ in their assumptions about how much information learners store and use. On the associative learning account, learners store and use a lot of data—all prior experiences gathered in the course of learning. On the hypothesis testing account, learners only use a small amount of data gathered on the most recent learning trial. To examine how much learning depends on the amount of prior experiences accumulated, we measured learning performance on each trial conditioned on the proportion of correct answers from all previous trials. We found that as total number of correct trials increases, performance on the current trial also increases proportionally ($\beta = 0.74$, $p < .001$; Fig. 5). For example, on the second trial, participants who gave a correct response on the first trial ($M = 0.58$, $SD = 0.49$, 95% CI [0.46, 0.69]) were more accurate than participants who responded incorrectly on the previous trial ($M = 0.24$, $SD = 0.43$, 95% CI [0.18, 0.29], $\beta = 1.19$, $p <.001$). Similarly, on the third trial, participants gave correct responses on both of the previous trials ($M = 0.82$, $SD = 0.39$, 95% CI [0.71, 0.93]) were more accurate than learners who were correct on only one of the two previous trials ($M = 0.48$, $SD = 0.50$, 95% CI [0.35, 0.60], $\beta = 1.10$, $p <.01$), or no previous trials ($M = 0.19$, $SD = 0.39$, 95% CI [0.13, 0.25], $\beta = 2.69$, $p <.001$).

To quantify this accumulated effect as a continuous variable, we added another factor to the mixed effects model—the total number of previous trials on which participants gave a correct response. An ANOVA showed that this new variable significantly improved model fit ($\chi^2 = 1139$, $p < .001$), suggesting that the total amount of accumulated information over time influences learning performance on the current trial; all previous experience matters to learning. A pragmatic reason to leverage all learning experiences is that newly acquired knowledge can be fragile, requiring repeated exposures to be consolidated in memory (Bion et al., 2013; Horst & Samuelson, 2008; Vlach & Johnson, 2013). The more the learners repeatedly make a correct response, the more likely they will stay on the correct answer.

3.2.3. Effect of non-immediate trials—

There are two plausible mechanisms through which previous trials can impact subsequent learning: “all-in-one” or “all-in-all.” On an “all-in-one” account, learners do not remember distant experiences. All previous experiences with a target word are condensed into the most recent hypothesis. On an “all-in-all” account, learners store and have access to not only the most recent experience but also distant past experiences. One way to distinguish “all-in-one” and “all-in-all” is to investigate learning performance followed incorrect responses. We compared two cases: (1) when all trials proceeding the immediate wrong trial were also incorrect and (2) when at least one of the trials before the incorrect previous trial was correct. As shown in Fig. 6, when at least one of the non-immediate pervious trials was right, accuracy on the current trial ($M = 0.40$, $SD = 0.49$, 95% CI [0.30, 0.52]) was significantly higher compared with cases in which all non-immediate previous trials were incorrect ($M = 0.13$, $SD = 0.34$, 95% CI [0.10, 0.16], $\beta = 0.76$, $p < .001$). Thus, even when learners guessed incorrectly on the most recent trial, they retained information from the trials before that one and leveraged this information in subsequent learning. Thus, learners not only integrated cross-situational statistics across consecutive learning situations but also across distant learning situations.

We also counted the number of correct trials proceeding the immediate previous trial and examined whether those past trials could influence accuracy on the current trial. If learners only rely on the immediately previous trial to either accept or reject a mapping, there should be no impact from trials before the immediately previous trials. As shown in Fig. 7, regardless of accuracy of the immediately previous trial, the more previous trials on which learners gave correct responses, the more likely they were to be correct on the current trial ($\beta = 0.79$, $p < .001$). This finding provides further evidence that statistical learners use information not only from the immediately previous learning trial, but also encode and use all their past learning experiences in ambiguous learning contexts.

In summary, Experiment 2 showed that learners integrate information across a sequence of highly ambiguous situations to find the correct referent. Moreover, the set of detailed analyses on trial-by-trial data provided converging evidence that the amount of information that learners carried over was related to real-time learning performance and that all past learning experiences matter to learning. Cross-situational learning is a cumulative and continuous process that involves tracking and aggregating past experiences.

4. Experiment 3

In Experiment 2, participants were tested immediately after each vignette in order to collect trial-by-trial responses to estimate the course of cross-situational learning. However, language learners in the real world are not explicitly asked to retrieve information every time they hear a word. Because retrieval itself has been shown to improve learning (e.g., Karpicke & Roediger, 2008), in Experiment 3 we measured learning performance only at the end of the continuous learning session, without inserting intermediate test trials. If the overall learning results from Experiment 3 are similar to those from Experiment 2, then the computational mechanisms revealed by trial-by-trial results from Experiment 2 may be more likely to be at play in real-world continuous learning scenarios with no interruptions.

4.1. Method

4.1.1. Participants—Twenty-two Indiana University undergraduates (6 Males, M_{age} = 20.00 years, SD_{age} = 1.31 years) participated for course credits. None had participated in previous conditions or other similar experiments.

4.1.2. Materials—The same 60 ambiguous vignettes used in Experiment 2 were used. They were grouped into 12 blocks. Each block contained five ambiguous trials presented in the same order as in Experiment 2.

4.1.3. Procedure—The instructions were similar to those used in Experiment 2, except that we asked participants to guess only once at the end of each block.

4.2. Results and discussion

We first compared participants' final response accuracy with the fifth trial's baseline accuracy tested in Experiment 1. After watching five vignettes all referring to the same target, participants' accuracy ($M = 0.50$, 95% CI [0.44, 0.57]) was significantly higher than the comparable trials' baseline accuracies ($M = 0.13$, 95% CI [0.08, 0.18], $\beta = 1.95$, $p <$ .001). This result demonstrated that learners were able to learn from ambiguous learning trials without being tested explicitly.

We then compared final accuracies in Experiment 3 with accuracies from the final trial of blocks in Experiment 2 ($M_{\text{exp2}} = 0.49$, 95% CI [0.43, 0.56]). We found no difference between Experiments 2 and 3 ($\beta = 0.08$, $p = .76$), suggesting that trial-by-trial testing did not impact learners' overall learning performance observed in Experiment 2.

5. Discussion

In three Human Simulation experiments using videos collected from the child's own view, we analyzed the course of cross-situational learning in order to determine which mechanisms people use to learn across multiple ambiguous learning situations. Specifically, we found evidence supporting that: (1) Naming events that occur during naturalistic parent–child joint play vary in their degrees of ambiguity. There is a subset of highly ambiguous naming instances, creating a word learning challenge. (2) Learners store and keep track of past knowledge when learning new words from highly ambiguous situations. Learners' performance increases as more information comes in, regardless of its quality, suggesting that cross-situational learning is a process that benefits from continuous information integration. (3) Learning performances is robust across different testing procedures.

5.1. Cross-Situational word learning is a continuous process, not one-and-done

Many studies of early word learning start with the referential uncertainty problem. Decades of research shows that the referential uncertainty can be solved through either one-shot learning from a single unambiguous learning situation (e.g., Behrend et al., 2001; Carey, 2010; Carey & Bartlett, 1978; Goodman et al., 1998; Heibeck & Markman, 1987; Horst & Samuelson, 2008; Jaswal & Markman, 2001; Markson & Bloom, 1997; Spiegel & Halberda, 2011; Waxman & Booth, 2000; Wilkinson & Mazzitelli, 2003; Woodward et al., 1994) or cross-situational learning from multiple ambiguous situations (Chen et al., 2018; Fitneva & Christiansen, 2011; Koehne & Crocker, 2015; Monaghan et al., 2015; Onnis et al., 2011; Wang & Mintz, 2018; Yu & Smith, 2007, Zettersten & Saffran, 2019). One-shot learning relies on clear learning situations in order to build correct word-referent mappings. Observational studies in the real world have documented that children and parents do sometimes jointly create clear learning situations in their daily life (Yoshida & Smith, 2008; Yu et al., 2009; Yu & Smith, 2012b). Moreover, experimental studies in the laboratory show that children are capable of using various cues to disambiguate a learning situation and build a correct word-object mapping in a single encounter (Baldwin, 1991, 1993; Tomasello & Farrar, 1986). Both social cues provided by parents and knowledge previously acquired by children seem to play a critical role in reducing the uncertainty (Frank et al., 2009; Goodman et al., 1998). However, the results of our study suggest that learners may adjust the mappings over the course of subsequent learning even after they acquire the mappings through "one-shot" learning. Even after making a correct response, learning was not done; learners did not always stay with the correct mapping on subsequent trials. Instead, they sometimes switched to an incorrect mapping and then switched back to the correct one later on. The dynamics of their trial-by-trial responses between right and incorrect answers suggest that learners constantly update their knowledge as new information enters the learning system. Learning is not one-and-done, but rather a continuous process. It involves not only discovering new knowledge from the sea of data but also consolidating newly acquired knowledge.

Recent research on memory development shows that retaining newly acquired knowledge is challenging for young children (Vlach, 2019). For example, even though the ability to fast-map a word to its correct referent emerges as young as 17 months of age (Halberda, 2003), 24-month-old infants do not retain newly learned word-referent mappings after a 5-min delay (Bion et al., 2013; Horst & Samuelson, 2008). Similar results have been found in 3-year-old children and adults (Vlach & Sandhofer, 2012). These results challenge the one-shot learning solution (Carey & Bartlett, 1978) because information retention may not be as easy as previously believed, and it is necessary to integrate memory constraints into any word learning account (Soh & Yang, 2021). Word learning may have to build on continuously accumulating statistical evidence from repeated exposures (Bion et al., 2013).

A continuous learning process built on statistical evidence over time may seem to be slow and inefficient relative to one-shot learning. However, because the learning environment is noisy, knowledge acquired on the first encounter with a word may be inaccurate or incomplete. Therefore, one-shot learning without prior knowledge can be fragile and error-prone. A more robust and reliable solution is for learners to continuously update their knowledge based on new information, perhaps changing less as stored representations converge. Especially in an early stage, a learning system can benefit by accumulating statistical evidence as much as possible to build a solid foundation. After this initial stage, one-shot learning becomes more robust and emerges as an efficient way to accelerate the speed of vocabulary acquisition. Several computational simulations (McMurray, 2007; Yu, 2008) have shown the computational mechanism through which the same associative learning model becomes much more efficient just by leveraging accumulated knowledge over time.

5.2. An integrated view of associative learning and hypothesis testing

What learning mechanism is responsible for the continuous learning process proposed above? Two prominent accounts of statistical learning—associative learning and hypothesis testing—have been traditionally characterized as fundamentally different. Recent computational and experimental evidence suggests the opposite. Rather than attempting to adjudicate between these two learning frameworks even more, a few recent behavioral studies (e.g., Romberg & Yu, 2014; Roembke & McMurray, 2016) were designed to investigate the potential interactions between explicit hypothesis testing and implicit associative learning processes. In a study done by Romberg and Yu (2014) using a cross-situational word learning paradigm, participants were asked to generate explicit hypotheses while aggregating trial-by-trial statistical information. They found that hypotheses are generalized based on the co-occurrence statistics accumulated through associative learning. This result suggests that both computational mechanisms may be involved in cross-situational learning (Romberg & Yu, 2014) and that they may work together in an integrated learning system. Relatedly, a recent variant of the PbV model called Pursuit (Stevens et al., 2017; Yang, 2020) has also demonstrated how hypothesis testing and associative learning can be integrated. Unlike PbV, in the Pursuit model, rejected hypotheses are not completely disregarded, but instead are retained for later evaluation. Similar to PbV, the Pursuit model also only stores one referent per learning instance. Stevens et al. (2017) ran a series of model comparisons and found that the Pursuit model outperforms associative learning and PbV models (Yang, 2020). Although the Pursuit model is still called a localist model where learners store one referent at the time, the fact that rejected hypotheses can be later retrieved and evaluated suggest that past knowledge plays a role and learning does not only rely on information provided at the moment.

In addition, computational analyses of the two learning mechanisms reveal that hypothesis testing can be viewed as a special case of associative learning when a simulated associative learner is selective and focus only on a small set of associations (Yu & Smith, 2012a). Similarly, Yurovsky and Frank (2015)argued that when learning a new word, participants not only encode the hypothesized referent but also encode several additional mappings at the same time and the different learning patterns found in support of either learning model depend on the complexity of the learning tasks.

5.3. Quantifying egocentric input using the HSP

HSP was originally developed to use adults to “simulate” child learners. Because adult learners have a developed conceptual system, using adults allows researchers to bypass concept development and isolate the problem of mapping those concepts to linguistic labels (e.g., Bloom, 2000; Gleitman, 1990; Smith, 2000). In the HSP, adults are instructed to guess what parents would say to their children by watching a third-person-view vignette. Because an event shown in third-person view can be highly ambiguous, adults often need to make an inference about what the parent may have in mind in order to guess what the parent may say at the moment. In the present study, however, egocentric videos from the child’s view reflect the learner’s personal view of an event, which have been found to be visually salient and less ambiguous (Yu & Smith, 2012b; Yurovsky et al., 2013). After watching a first-person view vignette, adult learners tend to rely on the perceptual information in view to make their guesses. In other words, instead of guessing what parents would say to their children in a particular context, adult learners are describing what they see using a linguistic label. Although this label may not be the one that children hear from parents, it reflects what perceptual information is embedded in the scene, therefore, can help us quantify the visual input children perceive in the moment. This measure is particularly useful for understanding word learning because to successfully learn a word, children need to derive word-to-world relations from the learning input. We can understand one important aspect of this problem by quantifying the alignment between what learners see with what they hear. HSP can be viewed as a good way to reveal what children see, as expressed in linguistic labels from adults.

Using this approach, a recent study provided new evidence on the inherent difficulty in learning verbs. For example, after seeing a vignette in which a parent is putting a toy phone close to her ear, adults provided multiple labels to describe the observed event, such as "putting," "calling," "talking," and "answering." All those linguistic labels could fit well in the situation, but only one of them may match with exactly what the parent said at the moment (Zhang et al., 2020). Inferring the correct word meaning from multiple equally suitable options is a challenging task for young learners. The linguistic labels provided by adults in the HSP can be directly compared with the words produced by parents at the same moments, which is a useful way to quantify the uncertainty that children experience from their own point of view.

6. Conclusion

Adapting the HSP, this research illustrates the value of characterizing the quality of naming events in the corpus of naturalistic interactions collected from infants' egocentric views. It also shows how we can understand the mechanisms responsible for integrating across the kind of ambiguous naming events that children see. Our results provide evidence that even though individual learning moments can be highly ambiguous, the statistical regularities embedded across multiple ambiguous naming events can still support learning by information integration. In other words, the correct signal can be found among considerable noise.

Acknowledgments

This research was supported by NICHD R01HD074601 and R01HD093792. We would like to thank Dr. Linda Smith, Dr. Sumarga Suanda, Dr. Chi-hsin Chen, Dr. Tian Xu, Dr. Catalina Suarez-Rivera, Seth Foster, Charlene Tay, and Steven Elmlinger for providing constructive feedback and Stella Huang for helping 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 DA (1991). Infants' contribution to the achievement of joint reference. Child Development, 62, 874–890.

Baldwin DA (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. Developmental Psychology, 29, 832.

Behrend DA, Scofield J, & Kleinknecht EE (2001). Beyond fast mapping: Young children's extensions of novel words and novel facts. Developmental Psychology, 37, 698. [PubMed: 11552764]

Bion RA, Borovsky A, & Fernald A (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. Cognition, 126, 39–53. [PubMed: 23063233]

Bloom P (2000). How children learn the meanings of words. Cambridge, MA: MIT Press.

Blythe RA, Smith K, & Smith AD (2010). Learning times for large lexicons through cross-situational learning. Cognitive Science, 34, 620–642. [PubMed: 21564227]

Carey S, & Bartlett E (1978). Acquiring a single new word. Proceedings of the Stanford Child Language Conference, 15, 17–29.

Carey S (2010). Beyond fast mapping. Language Learning and Development, 6, 184–205. [PubMed: 21625404]

Chen CH, Zhang Y, & Yu C (2018). Learning object names at different hierarchical levels using cross-situational statistics. Cognitive Science, 42, 591–605. [PubMed: 28685848]

Fitneva SA, & Christiansen MH (2011). Looking in the wrong direction correlates with more accurate word learning. Cognitive Science, 35, 367–380. [PubMed: 21429004]

Frank M, Goodman N, & Tenenbaum J (2009). Using speakers' referential intentions to model early cross-situational word learning. Psychological Science, 20, 578–585. [PubMed: 19389131]

Gillette J, Gleitman H, Gleitman L, & Lederer A (1999). Human simulations of vocabulary learning. Cognition, 73, 135–176. [PubMed: 10580161]

Gleitman L (1990). The structural sources of verb meanings. Language Acquisition, 1, 1–55.

Golinkoff RM, Hirsh-Pasek K, Bailey LM, & Wenger NR (1992). Young children and adults use lexical principles to learn new nouns. Developmental Psychology, 28, 99–108.

Goodman JC, McDonough L, & Brown NB (1998). The role of semantic context and memory in the acquisition of novel nouns. Child Development, 69, 1330–1344. [PubMed: 9839419]

Halberda J (2003). The development of a word-learning strategy. Cognition, 87, B23–B34. [PubMed: 12499109]

Heibeck TH, & Markman EM (1987). Word learning in children: An examination of fast mapping. Child Development, 1021–1034. [PubMed: 3608655]

Horst JS, & Samuelson LK (2008). Fast mapping but poor retention by 24-month-old infants. Infancy, 13, 128–157. [PubMed: 33412722]

Jaswal VK, & Markman EM (2001). Learning proper and common names in inferential versus ostensive contexts. Child Development, 72, 768–786. [PubMed: 11405581]

Kachergis G, Yu C, & Shiffrin RM (2012). An associative model of adaptive inference for learning word–referent mappings. Psychonomic Bulletin & Review, 19, 317–324. [PubMed: 22215466]

Karpicke JD, & Roediger HL (2008). The critical importance of retrieval for learning. Science, 319, 966–968. [PubMed: 18276894]

Koehne J, & Crocker MW (2015). The interplay of cross-situational word learning and sentence-level constraints. Cognitive Science, 39, 849–889. [PubMed: 25244041]

Macnamara J (1972). Cognitive basis of language learning in infants. Psychological Review, 79, 1–13. [PubMed: 5008128]

Markman EM, & Wachtel GF (1988). Children's use of mutual exclusivity to constrain the meanings of words. Cognitive psychology, 20, 121–157. [PubMed: 3365937]

Markman EM (1989). Categorization and naming in children: Problems of induction. Cambridge, MA: MIT Press.

Markson L, & Bloom P (1997). Evidence against a dedicated system for word learning in children. Nature, 385, 813–815. [PubMed: 9039912]

McMurray B (2007). Defusing the childhood vocabulary explosion. Science, 317, 631. [PubMed: 17673655]

Medina TN, Snedeker J, Trueswell JC, & Gleitman LR (2011). How words can and cannot be learned by observation. Proceedings of the National Academy of Sciences, 108, 9014–9019.

Monaghan P, Mattock K, Davies RA, & Smith AC (2015). Gavagai is as gavagai does: Learning nouns and verbs from cross-situational statistics. Cognitive Science, 39, 1099–1112. [PubMed: 25327892]

Onnis L, Edelman S, & Waterfall H (2011). Local statistical learning under cross-situational uncertainty. In Carlson L, Holscher C, & Shipley T (Eds.), Proceedings of the 33rd annual meeting of the Cognitive Science Society conference (Vol. 33, pp. 2697–2702). UC Merced, Merced, CA: eScholarship.

Quine WVO (1960). Word and object. Studies in Communication. New York, NY: Technology Press of MIT.

Rebuschat P, Monaghan P, & Schoetensack C (2021). Learning vocabulary and grammar from cross-situational statistics. Cognition, 206, 104475. [PubMed: 33220942]

Roediger HL III, & Butler AC (2011). The critical role of retrieval practice in long-term retention. Trends in Cognitive Sciences, 15, 20–27. [PubMed: 20951630]

Regier T (2005). The emergence of words: Attentional learning in form and meaning. Cognitive Science, 29, 819–865. [PubMed: 21702796]

Roembke TC, & McMurray B (2016). Observational word learning: Beyond propose-but-verify and associative bean counting. Journal of Memory and Language, 87, 105–127. [PubMed: 26858510]

Romberg A, & Yu C (2014). Interactions between statistical aggregation and hypothesis testing mechanisms during word learning. In Proceedings of the annual meeting of the Cognitive Science Society (Vol. 36, pp. 1311–1316). UC Merced, Merced, CA: eScholarship.

Scott RM, & Fisher C (2011). 2.5-Year-olds use cross-situational consistency to learn verbs under referential uncertainty. Cognition, 122, 163–180. [PubMed: 22104489]

Siskind J (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. Cognition, 61, 39–91. [PubMed: 8990968]

Smith K, Smith AD, & Blythe RA (2011). Cross-situational learning: An experimental study of word-learning mechanisms. Cognitive Science, 35, 480–498.

Smith LB (2000). How to learn words: An Associative Crane. In Golinkoff R & Hirsh-Pasek K (Eds.), Breaking the word learning barrier (pp. 51–80): Oxford, England: Oxford University Press.

Smith LB, & Yu C (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. Cognition, 106, 1558–1568. [PubMed: 17692305]

Smith LB, & Yu C (2013). Visual attention is not enough: Individual differences in statistical word-referent learning in infants. Language Learning and Development, 9, 25–49.

Soh C & Yang C (2021). Memory constraints on word learning. In Proceedings of the Annual Meeting of the Cognitive Science Society.

Spiegel C, & Halberda J (2011). Rapid fast-mapping abilities in 2-year-olds. Journal of Experimental Child Psychology, 109, 132–140. [PubMed: 21145067]

Stevens JS, Gleitman LR, Trueswell JC, & Yang C (2017). The pursuit of word meanings. Cognitive Science, 41, 638–676. [PubMed: 27666335]

Suanda SH, Mugwanya N, & Namy LL (2014). Cross-situational statistical word learning in young children. Journal of Experimental Child Psychology, 126, 395–411. [PubMed: 25015421]

Suanda SH, & Namy LL (2012). Detailed behavioral analysis as a window into cross-situational word learning. Cognitive Science, 36, 545–559. [PubMed: 22257004]

Thiessen ED (2017). What's statistical about learning? Insights from modelling statistical learning as a set of memory processes. Philosophical Transactions of the Royal Society B: Biological Sciences, 372, 20160056.

Tomasello M, & Farrar MJ (1986). Joint attention and early language. Child Development, 1454–1463. [PubMed: 3802971]

Trueswell JC, Medina TN, Hafri A, & Gleitman LR (2013). Propose but verify: Fast mapping meets cross-situational word learning. Cognitive Psychology, 66, 126–156. [PubMed: 23142693]

Vlach HA (2019). Learning to remember words: Memory constraints as double-edged sword mechanisms of language development. Child Development Perspectives, 13, 159–165.

Vlach HA, & Sandhofer CM (2012). Fast mapping across time: Memory processes support children's retention of learned words. Frontiers in Psychology, 3, 46. [PubMed: 22375132]

Vlach HA, & Johnson SP (2013). Memory constraints on infants' cross-situational statistical learning. Cognition, 127, 375–382. [PubMed: 23545387]

Vouloumanos A (2008). Fine-grained sensitivity to statistical information in adult word learning. Cognition, 107, 729–742. [PubMed: 17950721]

Vouloumanos A, & Werker JF (2009). Infants' learning of novel words in a stochastic environment. Developmental Psychology, 45, 1611–1617. [PubMed: 19899918]

Wang FH, & Mintz TH (2018). The role of reference in cross-situational word learning. Cognition, 170, 64–75. [PubMed: 28942355]

Waxman SR, & Booth AE (2000). Principles that are invoked in the acquisition of words, but not facts. Cognition, 77, B33–B43. [PubMed: 10986366]

Wilkinson KM, & Mazzitelli K (2003). The effect of "missing" information on children's retention of fast-mapped labels. Journal of Child Language, 30, 47–73. [PubMed: 12718293]

Woodward AL, Markman EM, & Fitzsimmons CM (1994). Rapid word learning in 13-and 18-month-olds. Developmental Psychology, 30, 553–566.

Yang C (2020). How to make the most out of very little. Topics in Cognitive Science, 12, 136–152. [PubMed: 30861339]

Yoshida H, & Smith LB (2008). What's in view for toddlers? Using a head camera to study visual experience. Infancy, 13, 229–248. [PubMed: 20585411]

Yu C, Ballard DH, & Aslin RN (2005). The role of embodied intention in early lexical acquisition. Cognitive Science: A Multidisciplinary Journal, 29, 961–1005.

Yu C, & Smith LB (2007). Rapid word learning under uncertainty via cross-situational statistics. Psychological Science, 18, 414–420. [PubMed: 17576281]

Yu C (2008). A statistical associative account of vocabulary growth in early word learning. Language Learning and Development, 4, 32–62.

Yu C, & Smith LB (2012a). Modeling cross-situational word-referent learning: Prior questions. Psychological Review, 119, 21–39. [PubMed: 22229490]

Yu C, & Smith LB (2012b). Embodied attention and word learning by toddlers. Cognition, 125, 244–262. [PubMed: 22878116]

Yu C, Smith LB, Shen H, Pereira AF, & Smith T (2009). Active information selection: Visual attention through the hands. IEEE Transactions on Autonomous Mental Development, 1, 141–151. [PubMed: 21031153]

Yurovsky D, Smith LB, & Yu C (2013). Statistical word learning at scale: The baby's view is better. Developmental Science, 16, 959–966. [PubMed: 24118720]

Yurovsky D, & Frank MC (2015). An integrative account of constraints on cross-situational learning. Cognition, 145, 53–62. [PubMed: 26302052]

Yurovsky D, Fricker DC, Yu C, & Smith LB (2014). The role of partial knowledge in statistical word learning. Psychonomic Bulletin & Review, 21, 1–22. [PubMed: 23702980]

Zettersten M, & Saffran JR (2019). Sampling to learn words: Adults and children sample words that reduce referential ambiguity. Developmental Science, 24, e13064.

Zhang Y, Amatuni A, Cain E, & Yu C (2020). Seeking meaning: Examining a cross-situational solution to learn action verbs using human simulation paradigm. In Proceedings of the Annual Meeting of the Cognitive Science Society (pp. 2854–2860). Wheat Ridge, CO: Cognitive Science Society.

Zhang Y, & Yu C (2017). How misleading cues influence referential uncertainty in statistical cross-situational learning. In LaMendola M & Scott J (Eds.), 41st annual Boston University conference on language development (pp. 820–833). Somerville, MA: Cascadilla Press.

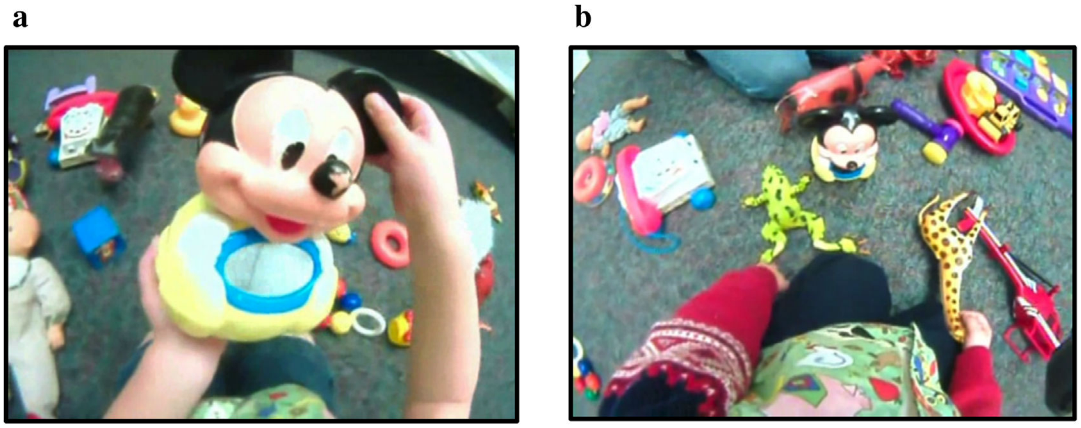

Fig. 1.

Both highly unambiguous (a) and highly ambiguous (b) vignettes were used in Experiment 1. The named object “mickey” can be easily identified in (a) as the dominant object in view, but not in (b) which contains multiple competing objects at the naming moment.

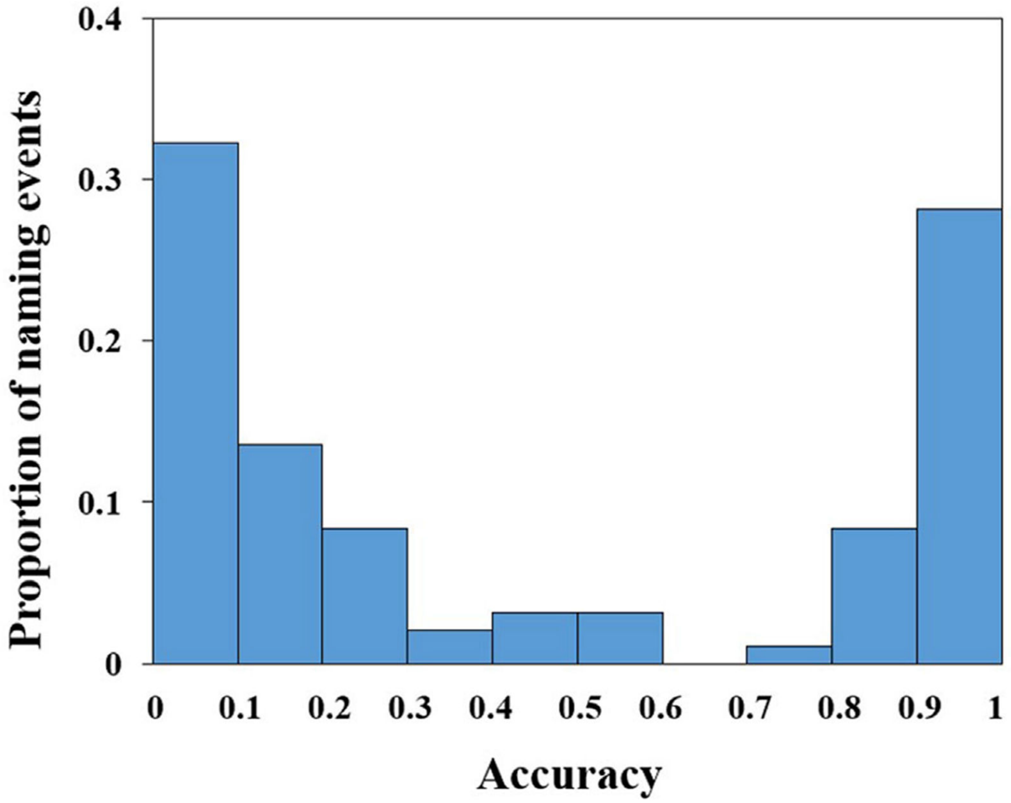

Fig. 2.

Distribution of response accuracy across vignettes. Consistent with Yurovsky et al. (2013), trial accuracies vary across different naming instances. Only ambiguous trials with less than 70% accuracy were used in the following studies.

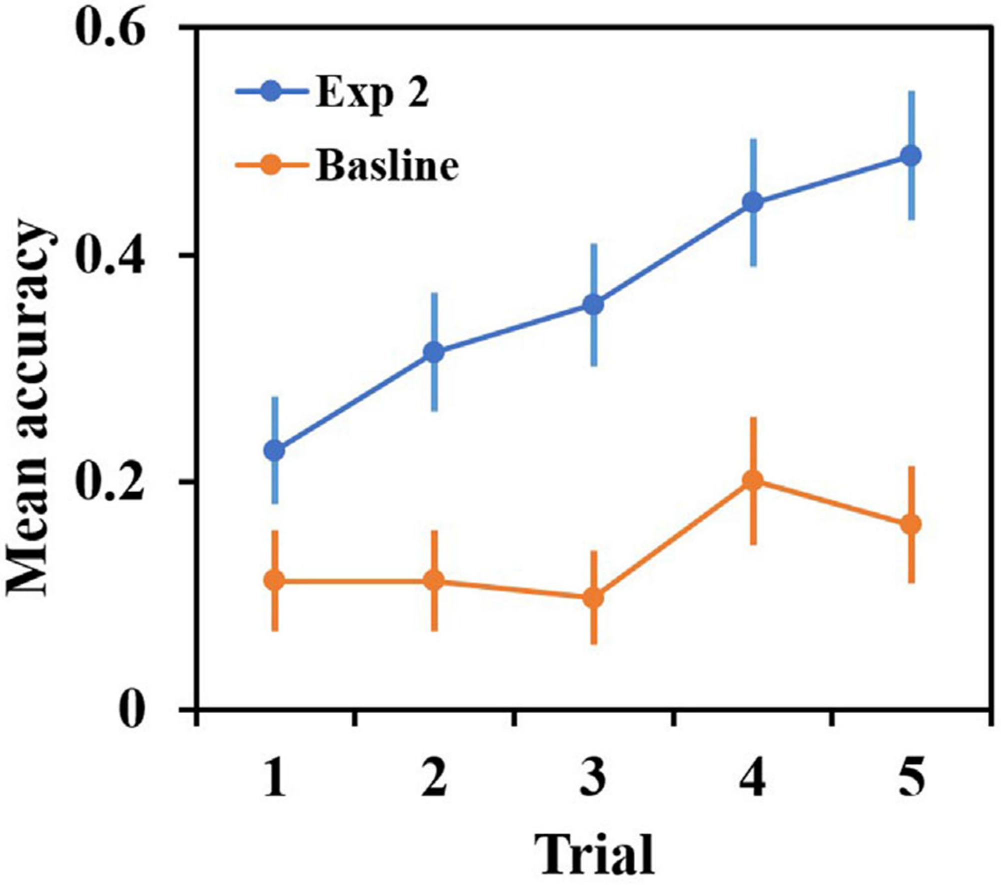

Fig. 3.

Mean accuracy across five ambiguous naming instances in Experiment 2 (Blue) and baseline accuracy of all ambiguous trials in Experiment 1 (Orange). Participants' response accuracy was significantly above baseline and improved across trials. Dots represent group means, and error bars represent 95% confidence intervals.

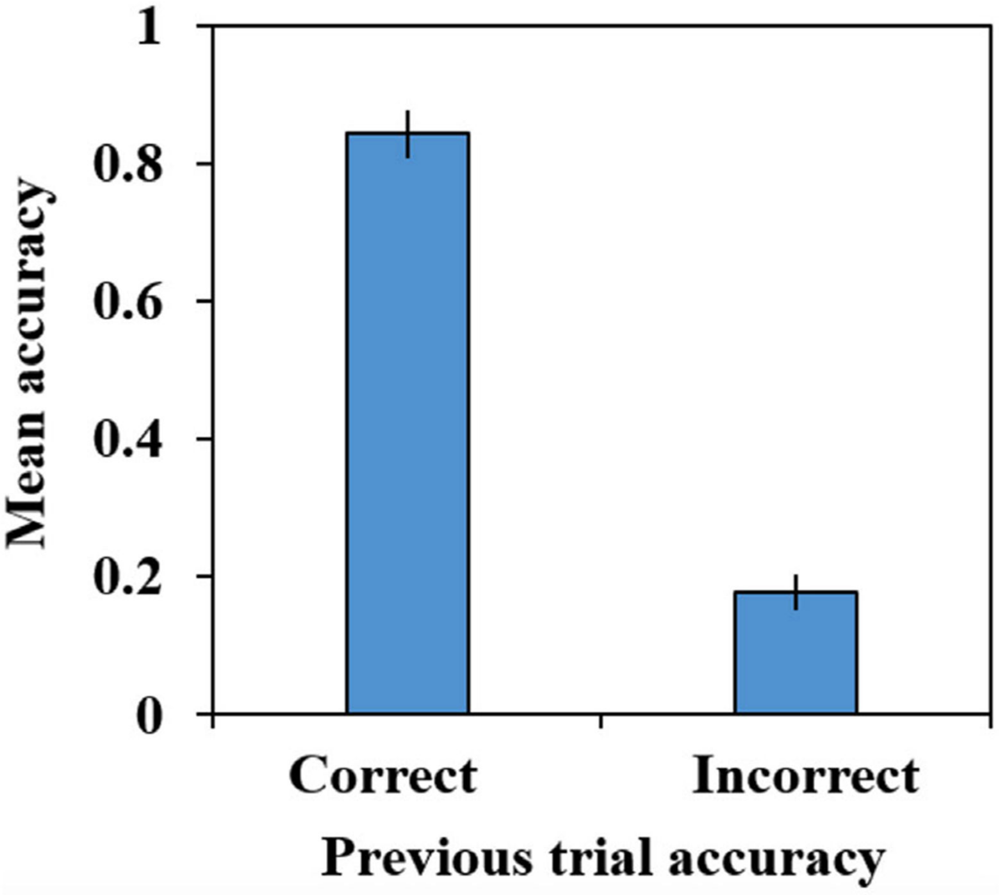

Fig. 4.
Current trial accuracy as a function of previous trial accuracy. Participants were more likely to respond correctly on the current trial if they also responded correctly on the previous trial. Bars represent group means, and error bars represent 95% confidence intervals.

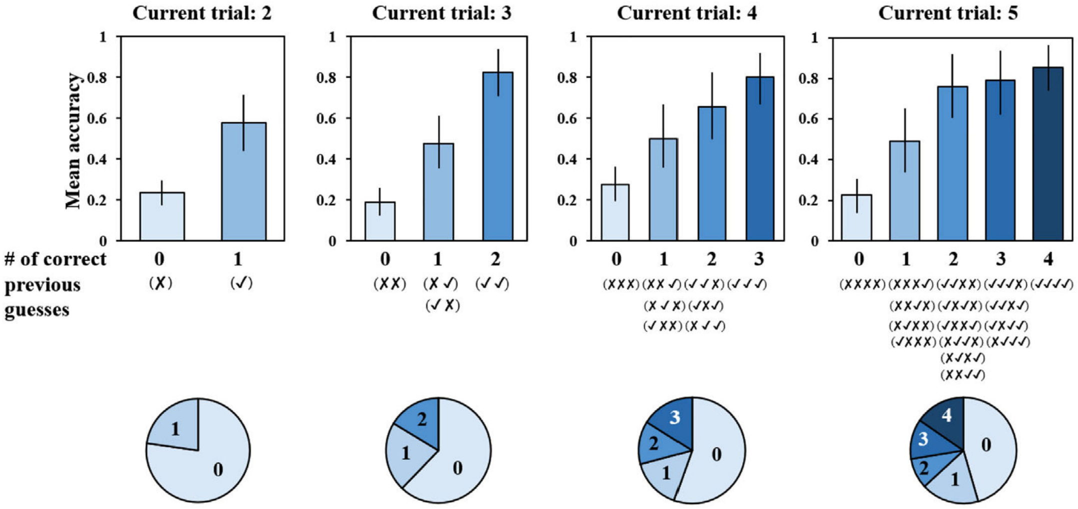

Fig. 5.

Accuracy on each trial as a function of the number of correct responses on previous trials. For each subplot, the *x*-axis shows the number of correct previous responses. Checks and crosses in the parentheses indicate all possible response patterns for previous trials. The *y*-axis shows current trial accuracy. Accuracy improved with more correct previous responses, indicating that the information accumulated on prior trials influences subsequent learning. Bars represent group means, and error bars represent 95% confidence intervals. Pie charts show the proportion of each type of instances.

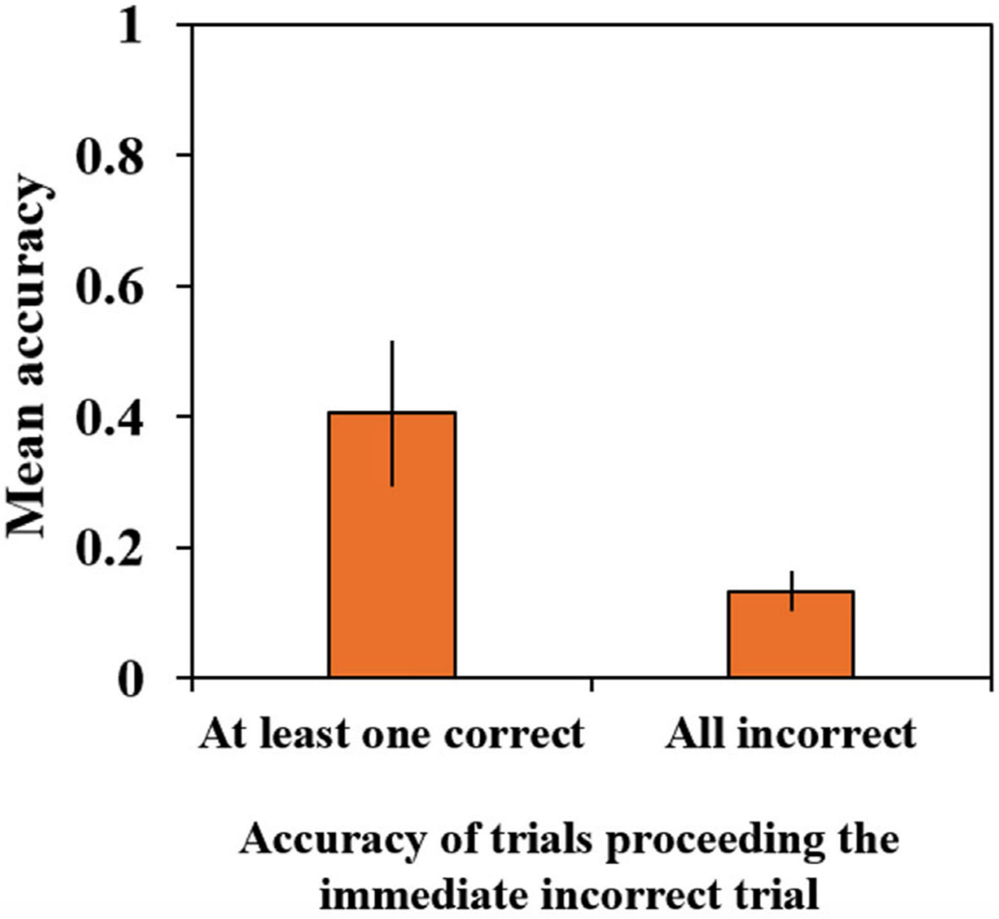

Fig. 6.

Among trials that were proceeded by an incorrect response, accuracies on the current trial differed depend on whether all non-immediately previous trials were incorrect or at least one of the non-immediately previous trials was correct. Bars represent group means, and error bars represent 95% confidence intervals.

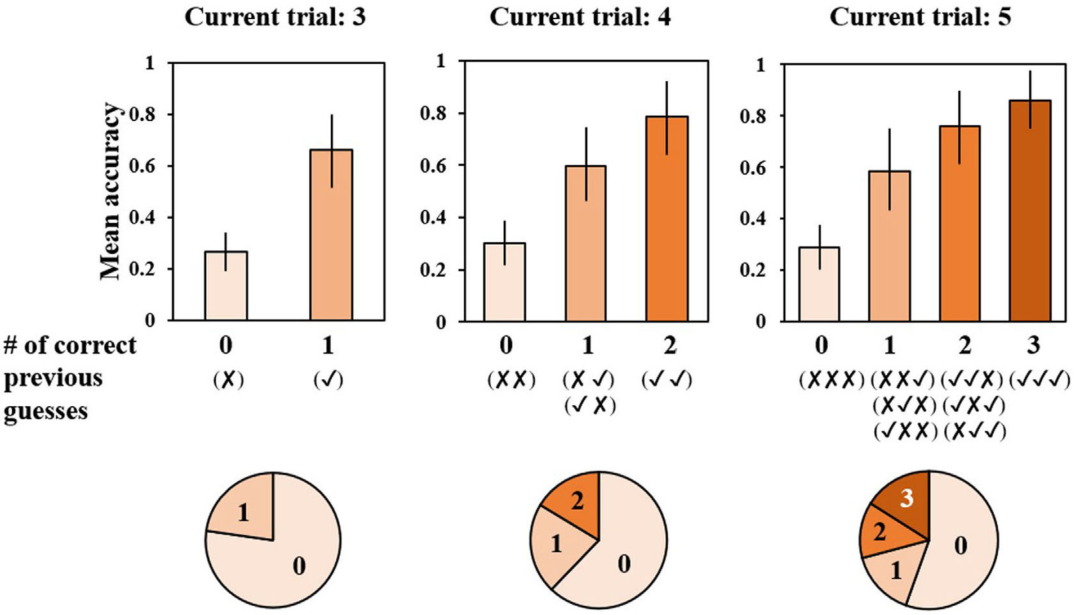

Fig. 7.

Accuracy (*y*-axis) as a function of the number of correct responses from *non-immediately* previous trials (*x*-axis). Regardless of accuracy of the immediate previously trial, accuracy on the current trial increased proportionally with the number of correct responses from non-immediately previous trials. Learners thus use information not just from the most recent previous trial but also from other more distant trials. Bars represent group means, and error bars represent 95% confidence intervals. Pie charts show the proportion of each type of instance.