Detailed Behavioral Analysis as a Window Into Cross-Situational Word Learning

Sumarga H. Suanda, Laura L. Namy

Department of Psychology, Emory University

Abstract

Recent research has demonstrated that word learners can determine word-referent mappings by tracking co-occurrences across multiple ambiguous naming events. The current study addresses the mechanisms underlying this capacity to learn words cross-situationally. This replication and extension of Yu and Smith (2007) investigates the factors influencing both successful cross-situational word learning and mis-mappings. Item analysis and error patterns revealed that the co-occurrence structure of the learning environment as well as the context of the testing environment jointly affected learning across observations. Learners also adopted an exclusion strategy, which contributed conjointly with statistical tracking to performance. Implications for our understanding of the processes underlying cross-situational word learning are discussed.

Keywords: Cross-situational learning; Word learning; Language acquisition

1. Introduction

The notion that children learn words at least in part by tracking the common contexts in which a given word is uttered is neither a novel nor an unintuitive idea. Although many language acquisition scholars (e.g., Carey, 1978; Gleitman, 1990; Pinker, 1989) have argued that cross-situational learning plays a pivotal role in children's lexical acquisition, empirical demonstrations of this learning capacity are relatively scarce. Indeed, research on how children learn the meanings of words has focused predominantly on the mechanisms underlying children's word learning within a single observation or naming event (so called *fast mapping*; for reviews see Bloom, 2000; Golinkoff et al., 2000).

Correspondence should be sent to Sumarga H. Suanda, Department of Psychology, Emory University, Atlanta, GA 30322. E-mail: ssuanda@emory.edu

Recently, however, a number of studies have revealed that infants (Smith & Yu, 2008), toddlers (Akhtar & Montague, 1999), and adults (Gillette, Gleitman, Gleitman, & Lederer, 1999; Yu & Smith, 2007) are capable of learning words across multiple observations. For example, in one recent study, Yu and Smith (2007) presented adult learners with a series of naming events in which they heard multiple words and saw multiple pictures. Although each word always occurred with its referent across naming events, other words and other pictures were also present during any given event, leading to referential ambiguity within each event. Importantly, words and pictures were repeated in various permutations across naming events, leading to more reliable co-occurrences between words and their referents than between words and distractor objects. Despite only viewing a handful of repetitions of each word-picture pairing, adult learners correctly identified which words co-occurred most reliably with which pictures at above chance rates (Yu & Smith, 2007).

Yu and Smith's findings have more recently been extended to infant word learners (Smith & Yu, 2008; Yu & Smith, 2011; see also Vouloumanos & Werker, 2009), highlighting the fact that this learning capacity is readily available early in development and thus raising the possibility that it plays a role in children's early word learning. Despite this recent evidence of cross-situational word learning across development, the processes and factors that make this form of learning possible are relatively unknown. The broad goal of the current study is to shed light on the mechanisms underlying cross-situational word learning.

The logic behind successful learning in Yu and Smith's task, and cross-situational word learning more generally, is that within-trial or within-situation referential ambiguity requires learners to track multiple possible referents for each observed use of a word. Learners can then compare the set of possible referents across observations to arrive at the most probable referent for each word (Siskind, 1996; Yu & Smith, 2007). Recently, K. Smith and colleagues have proposed that a process other than computation of cross-situational statistics may also account for Yu and Smith's findings (Smith, Smith, & Blythe, 2009). They argue that a learner who simply keeps track of the set of objects and words present during a single learning trial could successfully perform above chance in Yu and Smith's task. This is possible because the testing regimen implemented by Yu and Smith (a four-alternative forced-choice task) constrains learners' referent selections to four items (the target picture and three foil pictures). Importantly, the probability with which the three foil items also co-occurred with the target word during the single encoded learning trial was relatively low. Thus, during test, the learner could perform at above-chance rates by simply selecting at random from the available pictures that had also been presented during the single encoded learning trial for that particular word.

Smith et al.'s (2009) alternative account highlights the notion that a variety of mechanisms could potentially explain successful cross-situational word learning. In the current experiment, we investigate these mechanisms by replicating Yu and Smith's original finding and examining the factors influencing successful learning as well as mis-mappings. We propose to accomplish this in the following ways. First, we examine the role of the statistical structure of the learning environment (as stressed by Yu and Smith) on the learning process. As described above, within the learning phase of Yu and Smith's cross-situational

word learning paradigm, learners view a series of ambiguous situations involving multiple words and multiple pictures. Thus, a cross-situational statistical learner creates associations not only between words and their correct referents but also between words and distractor pictures that are the referents of other words also presented during the trial (*spurious correlations*). As learning trials are constructed by randomly selecting four word-referent pairs for each trial, there is variability in the strength and number of possible spurious correlations formed for each word across trials. That is, during learning, some words may appear many times with few distractors, creating a small number of strong spurious correlations. In contrast, other words may appear few times with many distractors, creating many weak spurious correlations. In the current experiment, we ask to what extent these variations (i.e., differences in *contextual diversity*, Kachergis, Yu, & Shiffrin, 2009) affect learning.

Three recent findings suggest that the structure of the learning environment does have an influence on statistical word learning. First, in Yu and Smith's original study, participants in a condition with a larger to-be-learned lexicon acquired more words than participants in a condition with a smaller lexicon to learn. Yu and Smith argued that this pattern reflects the fact that the smaller lexicon resulted in fewer competitors and thus stronger spurious correlations. Second, Kachergis et al. (2009) demonstrated that systematically manipulating the diversity in the learning contexts in which words appear impacts cross-situational statistical learning; within the same to-be-learned lexicon, words with many weak spurious correlations elicited higher learning rates during test than those with fewer, stronger spurious correlations. A third reason to suspect that spurious correlations may affect learning is that word learners appear to readily track multiple word-to-referent mappings (Vouloumanos, 2008; Vouloumanos & Werker, 2009). That is, adult word learners (and to some extent infant word learners, see Vouloumanos & Werker, 2009) appeared sensitive not only to high-frequency pairings but also to low-frequency pairings. Based on these findings, we predict that variability in the co-occurrence structure of the learning environment should affect which words are learned best.

A second goal of the current experiment is to examine whether the testing environment's context influences performance independent of the structure of the learning environment. Test trials in this design involved presenting a target word with four possible referents (the target picture and three foils). As foils were randomly selected, some words were tested with foils that co-occurred often with the target word during learning while other words were tested with foils that co-occurred rarely or not at all with the target word during learning. We predicted that performance should vary as a function of the probability with which foils served as distractors during learning. Thus, in addition to exerting influence within the learning process, co-occurrence statistics should also have an effect during test. Of particular interest is the extent to which these effects are independent of one another, as this may disambiguate between the single-exposure learner and statistical learner accounts of cross-situational word learning. We suggest that although both cross-situational learning and single-trial accounts would predict an effect of testing environments on performance, only the cross-situational learning account would predict an independent effect of the learning environment's statistical structure on performance.

A third goal of the current experiment is to examine influences of additional mechanisms beyond tracking of co-occurrence statistics on cross-situational word learning. One candidate mechanism that has previously been proposed to play a prominent role in cross-situational word learning (e.g., Siskind, 1996; Yu, 2008) is *mutual exclusivity* (Markman & Wachtel, 1988). Briefly, the mutual exclusivity constraint refers to a word learner's default tendency to accept only one word for each object (Markman, 1992). This assumption may aid cross-situational word learning in several ways. Across the learning phase, mutual exclusivity simplifies the learning process by limiting the hypothesis space and guiding the learner away from entertaining many-to-one or one-to-many word-referent mappings. Mutual exclusivity may also contribute to performance as learners can use known word-referent mappings to rule out possible referents for unknown words, either during learning (Ichinco, Frank, & Saxe, 2009; Yurovsky & Yu, 2008) or during test (e.g., Diesendruck & Markson, 2001; Markman & Wachtel, 1988). In this study, we investigate participants' use of this strategy in the cross-situational learning paradigm by examining the extent to which knowledge of (i.e., having successfully mapped labels for) the foils at test constrains referent selection for the target word.

A final goal of the current study is to provide some insight into the automatic and non-strategic nature of participants' learning via cross-situational observations. Yu and Smith (2007) reported anecdotal evidence that participants in their task verbally reported learning very few words (see also Ichinco et al., 2009). This anecdotal evidence is reminiscent of previous investigations demonstrating that the learning of linguistic and non-linguistic statistical structures proceeds incidentally (e.g., Saffran, Newport, Aslin, Tunick, & Barrueco, 1997) and automatically (e.g., Turke-Browne, Junge, & Scholl, 2005). As a first step in assessing the automaticity of cross-situational word learning, we compared participants' performance to their own explicit judgments of their performance during a post-experiment interview.

2. Methods

2.1. Participants and stimuli

Forty-five adults participated for course credit or cash compensation. Our design replicated the 4 × 4 condition of Yu and Smith's (2007) first experiment. We recorded 54 spoken bisyllabic novel words (e.g., ''*bicket*'') and generated 54 pictures of uncommon or novel objects (e.g., a phototube). Each word was randomly paired with a picture and these pairs were divided into three blocks of 18 word-picture pairings.

2.2. Design and procedure

In each block, participants completed a learning phase followed by a test phase. In each trial of the learning phase, participants saw four randomly selected pictures appearing simultaneously on a 17-inch computer monitor, one in each quadrant, and heard four spoken

Fig. 1. Example series of learning trials (A) and test trial (B). During learning in the experiment no word-picture pairing appeared on back-to-back learning trials.

words corresponding to the four pictures played sequentially in a randomized order (with 2 s of silence between words, see Fig. 1). Each of the 18 word-picture pairings occurred in six trials, yielding a total of 27 learning trials per block. Because participants heard four words on each learning trial, words co-occurred not only with the correct corresponding picture but also with other distractor pictures (accompanied by their corresponding words). The association matrix presented in Fig. 2 illustrates an example of the relative frequencies with which words (columns) co-occur with different pictures (rows) throughout learning.

Although all words occurred six times during the learning phase and appeared with three distractors on each learning trial, the number of *different* distractors with which a word co-occurred as well as the *frequency* with which a given distractor co-occurred with a given word varied across items due to the randomized construction of the learning trials. The frequency with which any given word appeared with any given distractor varied from 0 to 4 times ($M = 1.10$, $SD = 0.091$). The number of unique distractors with which any given word co-occurred varied from 8 to 15 ($M = 11.96$, $SD = 1.51$). Thus, some words co-occurred frequently with a smaller number of distractor pictures, creating the potential for learners to detect a small number of strong spurious correlations, whereas other words co-occurred infrequently with a larger number of distractors, creating the potential for many weak spurious correlations.

The test phase immediately followed the learning phase for each block and consisted of 18 four-alternative forced-choice test trials, one per target word. In each trial, four pictures appeared simultaneously followed after 2 s by the presentation of one word. Participants indicated using the mouse which picture went with the target word. Test trials were constructed by selecting the target word's corresponding picture and three randomly selected foils, with all pictures serving as foils equally often. Random selection of test foils yielded variability in the associative strength of foils for each target word. After completing all three blocks, participants estimated their accuracy, responding to the question, ''What percentage of words do you think you got right?''

		Words																	
		W1	W2	W3	W4	W5	W6	W7	W8	W9	W10	W11	W12	W13	W14	W15	W16	W17	W18
Pictures	P1	6	3	1	1	2	1	2	1	2			1		2				2
	P2	3	6	2	1	2	1			2			2		1	1	1	1	1
	P3	1	2	6	2		2	2	3			1	2		1		1	1	
	P4	1	1	2	6	1		2	2	1	1	1			1		2	2	1
	P5	2	2		1	6	1	2	1	1	1	1		2	2		1	1	
	P6	1	1	2		1	6		2		2		1	2	2	1	1	1	1
	P7	2		2	2	2		6	3	1	1	1	1	2					1
	P8	1		3	2	1	2	3	6	1		1	1		1		1		1
	P9	2	2		1	1		1	1	6		1	2		2	3	1		1
	P10				1	1	2	1			6	2	1	4		1	2	3	
	P11			1	1	1		1	1	1	2	6	2	2	2	1	1	1	1
	P12	1	2	2			1	1	1	2	1	2	6			2	1		2
	P13					2	2	2			4	2		6		2		3	1
	P14	2	1	1	1	2	2		1	2		2			6	2	1		1
	P15		1				1			3	1	1	2	2	2	6	1	1	3
	P16		1	1	2	1	1		1	1	2	1	1		1	1	6	3	1
	P17		1	1	2	1	1				3	1		3		1	3	6	1
	P18	2	1		1		1	1	1	1		1	2	1	1	3	1	1	6

Fig. 2. Sample association matrix of word and picture pairings. Figure indicates how many times each word co-occurred with each picture across learning trials for a single learning block. Each word occurred on six learning trials. Empty cells denote that the word did not co-occur at all with a given picture.

3. Results

We closely replicated the learning levels demonstrated by Yu and Smith's participants. Participants learned roughly half of the items ($M = 0.503$, $SD = 0.18$), a significantly higher proportion than predicted by chance (0.25), $t(44) = 9.55$, $p < .001$, $d = 1.42$. Interestingly, their performance estimates ($M = 0.34$, $SD = 0.22$) were significantly lower than their actual performance, $t(44) = 4.59$, $p < .001$, $d = 1.44$, consistent with previous anecdotal evidence that participants were unaware of their learning success. However, participants' performance estimates were significantly correlated with their actual performance, $r(43) = .363$, $p < .05$, suggesting some explicit awareness of performance (see Fig. 3).

3.1. Effects of distractor co-occurrence and test foil interference

We examined whether learners formed spurious correlations with high probability distractors during learning and the extent to which these spurious correlations interfered with learning. To test this we calculated, for each word, the average probability with which distractors co-occurred across the six learning trials in which the word occurred. Each word had 18 possible distractor slots (three per trial over six trials). Some candidate distractor pictures were presented more often than others and some not at all. To illustrate, in the
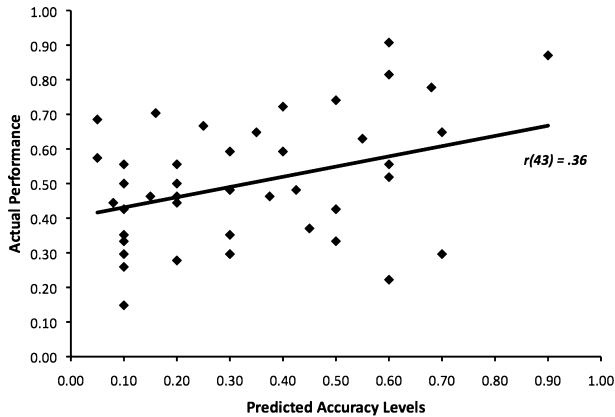
Fig. 3. Correlation between participants' estimated performance and actual performance levels.

example, in Fig. 2, Word 1 (W1) co-occurred with one distractor picture (Picture 2) on three of the six learning trials (resulting in .50 probability of co-occurrence), five pictures (Pictures 3, 4, 6, 8, 12) on two trials each (.333 probability) and five pictures (Pictures 5, 7, 9, 14, 18) on one trial each (.167 probability). Thus, average probability of distractor co-occurrence for this word was .27.

Average distractor co-occurrence probability varied from .20 to .375 ($M = 0.255$, $SD = 0.03$). For each participant, we compared the average distractor co-occurrence probability for the items in which subjects answered correctly versus incorrectly during test. Average distractor co-occurrence probability was slightly but significantly greater for incorrect ($M = 0.259$, $SD = 0.007$) than correct items ($M = 0.253$, $SD = 0.005$,), $t(44) = 3.98$, $p < .001$, suggesting that learners' detection of spurious correlations during learning interfered with the learning of that word.

To examine the effects of the relationship between distractors during learning and foils at test, we calculated average test foil strength for each word, based on the probability with which each of the three foils in the test trial had co-occurred with the target word during the learning phase. For example, if Pictures (P) 2, 8, and 10 (see Fig. 2) acted as test foils for W1, the test foil strength for W1 would be .22. This is obtained from averaging the probabilities with which P2 (.50), P8 (.167), and P10 (.000) had co-occurred with W1 during learning.

Average test foil strength ranged from 0 to .389 ($M = 0.179$, $SD = 0.08$). We compared the average test foil strength for items on which participants responded correctly versus incorrectly. We found that average test foil strength was higher for incorrect items ($M = 0.184$, $SD = 0.02$) than correct items ($M = 0.173$, $SD = 0.02$), $t(44) = 2.41$, $p < .05$, suggesting that the presence of foils in test that had co-occurred with the target word during learning and the frequency with which they had co-occurred influenced performance during test.

One goal of this experiment was to examine the relative contributions of distractor co-occurrence probability during learning and test foil strength during test in predicting participants' learning. An independent effect of distractor co-occurrence probability during

learning (i.e., contextual diversity) would support a cross-situational, as opposed to a single-trial learning account of participants' learning strategy. To test this, we employed a generalized linear mixed model with generalized estimating equations. This analysis was selected because the dependent variable was binary (correct vs. incorrect), as well as to account for the fact that each participant contributed multiple data points. The model predicted whether an item would be answered correctly as a function of (a) the item's average distractor co-occurrence probability, and (b) the item's average test foil strength.

The results of the model can be seen in Model 1 in Table 1. We found that the coefficients for both distractor co-occurrence probability, $p < .001$, and test foil strength, $p = .001$, were significant. These findings were consistent across blocks. A model that included block as a predictor, as well as interaction terms between block and our variables of interest, revealed no significant effect of block and no significant interactions. Thus, the structure of the learning environment and the relationship between learning and test environments exerted distinct influences on performance in this cross-situational learning paradigm. These findings indicate that learners cannot be employing exclusively a single-trial learning strategy.

These conclusions are bolstered by an analysis of individual patterns. We classified each participant's performance based on whether his or her performance was affected by (a) the statistical structure of the learning environment, (b) the associative strength of the foils in testing, (c) both, or (d) neither. Participants' performance was considered affected by the structure of the learning environment if their average distractor co-occurrence probability was greater for incorrect compared to correct items. Likewise, participants' performance was considered affected by the testing structure if their average test foil strength was greater for incorrect compared to correct items. As can be seen in Table 2, a large majority of participants were jointly affected by both the structure of the learning environment and testing context ($\chi^2 = 19.44$, $p < .001$), consistent with the group-level findings. Nonetheless, some participants appeared affected by only one of the two factors and some were not systematically influenced by either, underscoring that there is variability in the sources of information that learners reliably tracked.

Table 1
Coefficient estimates for mixed-model logistic regressions predicting item accuracy

	Statistics		
	Wald χ^2	Sig	Odds Ratio (OR)
Model 1			
Predictors			
Distractor co-occurrence probability	13.4	<.001	0.54
Test foil strength	10.12	.001	0.17
Model 2			
Predictors			
Distractor co-occurrence probability	11.48	.001	0.52
Test foil strength	8.03	.005	0.17
Known items	100.81	<.001	1.85

Table 2
Individual pattern distributions of factors affecting performance

		Test Foil Strength	
		INC > COR	INC ≤ COR
Distractor co-occurrence probability	INC > COR	24	7
	INC ≤ COR	6	8

Note. COR, correct items; INC, incorrect items.

3.2. Analysis of error patterns

We also examined learners' sensitivity to co-occurrence statistics by analyzing participants' error patterns. We reasoned that the probability with which an item would be selected erroneously during test should reflect the co-occurrence strength between that item and the target word established during learning. To test this, we calculated, for each participant, the proportion of trials on which participants selected foils at test that had co-occurred with target words on 0, 1, 2, 3, or 4 of the 6 learning trials (corresponding to co-occurrence probabilities of 0, .167, .333, .5, and .667, respectively). An analysis of variance (ANOVA) on proportion of choices revealed a significant effect of level of co-occurrence, $F(4, 44) = 5.78$, $p < .001$, $\eta^2_p = .116$. Follow-up comparisons revealed that there were no differences among the 0, .167, .333, and .5 probability foils (smallest $p > .10$). However, as seen in Fig. 4, participants selected the .667 probability foil significantly more often than they did the lower probability foils (all ps $< .05$). This finding highlights participants' sensitivity to word-distractor co-occurrence statistics but also suggests that participants were particularly lured by the presence of a high probability competitor. At first glance, these findings appear to contradict Yu and Smith's finding that there was no systematicity in foil selection based on foil-target co-occurrence. However, Yu and Smith's original design did not include
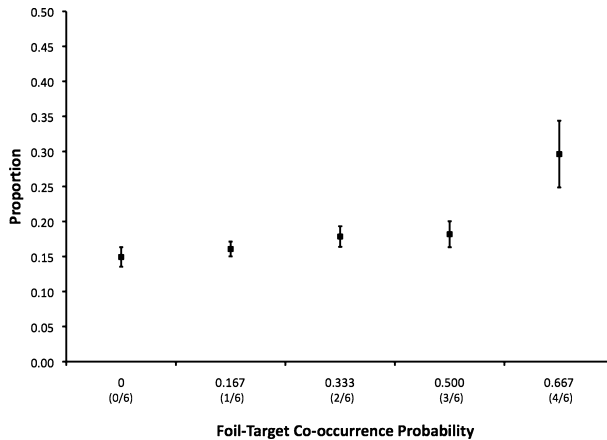

Fig. 4. Proportion of foil selection as a function of the target-foil co-occurrence probability.

any foils that co-occurred with the target with a probability above .50, raising the possibility that only foils with a relatively high threshold of co-occurrence may lure participants. Although performance was disproportionately affected by high co-occurring foils, this does not imply that lower levels of foil-target probability have no effect on behavior. When we repeated our models and excluded trials containing the highest probability competitors (i.e., the .667 probability foils), test foil strength ($ps < .01$) nonetheless remained a significant predictor.[1]

3.3. Exclusion constraint on task performance

We also discovered that task success was not driven exclusively by cross-situational statistics. Knowledge of the correct word mappings for foils (as indicated by participants' accuracy when those foils were tested as targets) also constrained performance on test trials. Specifically, accuracy on test trials varied as a function of the number of foil objects (between 0 and 3) participants had correctly mapped (see Fig. 5). An ANOVA comparing accuracy as a function of the number of foils known revealed a significant main effect of foil label knowledge on item accuracy, $F(3, 36) = 12.69$, $p < .001$, $\eta^2_p = .25$. Importantly, even for trials in which participants knew none of the foils, performance was still significantly above chance, $t(41) = 2.89$, $p < .01$.

We conducted a second mixed-model logistic regression that was identical to the model described above with the exception that number of known foils was added as a third predictor variable. As seen in Model 2 reported in Table 1 all three variables independently contributed to predicting item accuracy. These findings were consistent across blocks. A model that included block as a predictor yielded no significant effect of block, and no significant interactions. This suggests that statistical computations and elimination of foil candidates as potential referents conjointly contributed to performance in this cross-situational word learning paradigm.
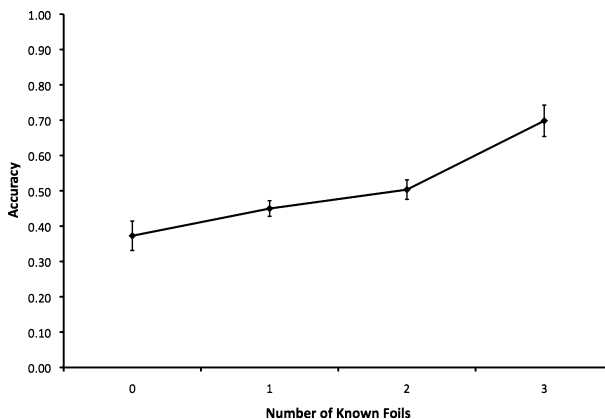

Fig. 5. Mean accuracy as a function of the number of known foils.

4. Discussion

A number of recent reports studying a wide range of age groups have convincingly demonstrated that word learning through the computation of cross-situational statistics is well within the repertoire of human learners (Akhtar & Montague, 1999; Gillette et al., 1999; Smith & Yu, 2008; Yu & Smith, 2007). Furthermore, several computational models have successfully learned words cross-situationally using inputs derived from transcriptions of real parent labeling behavior (Frank, Goodman, & Tenenbaum, 2009; Yu, 2008; Yu & Ballard, 2007). This highlights the plausible role of cross-situational learning as a mechanism underlying children's lexical acquisition. The goal of the current study was to use detailed behavioral analyses as a window into the range of mechanisms and factors influencing cross-situational word learning.

4.1. Effects of co-occurrence statistics during learning and test

Item and error analyses revealed that the strength of spurious correlations during learning and the associative strength between the target word and foils present at test independently affected cross-situational learning. We suggest that the independent effect of the structure of the learning environment on learning challenges the notion that learners' success can be accounted for exclusively by single-trial learning (Smth et al., 2009). That is, although both a single-trial learning and a cross-situational learning account predict an effect of the testing context on successful learning, only a cross-situational learning account predicts an independent effect of statistical structure of the learning environment on learning. This is because differing levels of spurious correlations (a measure of learning environment structure) arise only *across* observations, and a single-exposure learner would not encode multiple observations of a given word. Thus, the effect of spurious correlations during learning implies that participants tracked co-occurrence statistics across multiple learning events rather than simply encoding an individual naming event for each word. These results are consistent with Yu and Smith's original proposal of how learning in this task occurs across situations, as well as the finding that adult learners show sensitivity to fine-grained co-occurrence information in other word-learning paradigms (Vouloumanos, 2008).

There are two possible explanations for why spurious correlations during learning may have influenced performance. First, items with low spurious correlations may have been easier to learn because the words occurred across more variable contexts (i.e., with a greater diversity of distractors) during learning. Indeed, recent evidence suggests that greater contextual diversity aids cross-situational word learning (Kachergis et al., 2009). Alternatively, the items with high spurious correlations may have been more difficult to learn due to the presence of strong competitors (distractors that co-occurred often with a word). Although future work may be able to tease apart the effects of contextual diversity from that of strong foil competition, contextual diversity and strong competition are typically highly negatively correlated.

4.2. Constraints on cross-situational word learning

Although our findings lend support to the notion of statistical learning mechanism underlying cross-situational learning, we found evidence that processes other than statistical learning also played a role. Specifically, performance on this task varied as a function of participants' knowledge of foils during test. That is, participants were more likely to select the correct target if they had correctly mapped labels to the foil pictures present during the test trial. This pattern of result suggests the use of an exclusion strategy at test, identifying the correct referent by ruling out alternatives. This finding underscores the fact that in-the-moment mechanisms can augment cross-situational statistical word learning (see also, Ichinco et al., 2009; Siskind, 1996; Yu, 2008; Yu & Smith, 2007; Yurovsky & Yu, 2008). It is clear from these findings that *both* exclusion and statistical learning contributed to performance because (a) the effects of distractor co-occurrence probability and test foil strength on item accuracy were independent of the effect of exclusion; (b) even in situations where participants knew none of the foils, performance was still above chance rates; and (c) use of foil knowledge as a basis for exclusion, would have required acquisition of foil knowledge, at least in part, via cross-situational learning. Thus, our findings reflect how participants recruit statistical learning and word learning constraints conjointly to determine word-referent mappings.

4.3. The automatic nature of cross-situational word learning

Our finding that participants' explicit judgments vastly underestimated their actual learning rates is consistent with previous anecdotal evidence of participants' lack of awareness of cross-situational word learning (Ichinco et al., 2009; Yu & Smith, 2007). Given the coarseness of our measure of participants' awareness of learning, future studies should utilize more established indices of implicit knowledge, such as associative priming (see Seger, 1998), to provide a more direct test of the implicit nature of cross-situational learning. However, despite its simplicity, participants' verbal reports were positively correlated with their learning rates, suggesting that our explicit measure captured some sensitivity to learning.

4.4. Implications and future directions

One important avenue by which researchers have gained insight into the mechanisms underlying cross-situational learning is through computational modeling (e.g., Frank et al., 2009; Siskind, 1996; Yu, 2008). However, as some have noted (Frank et al., 2009; Ichinco et al., 2009; Yu, Smith, Klein, & Shiffrin, 2007), a number of computational instantiations using distinct underlying architectures and assumptions can successfully model accuracy in Yu and Smith's cross-situational learning task, leaving open the question of which model best approximates learners' performance. We suggest that comparing how different models account for patterns of behavior beyond accuracy (such as effects of spurious correlations during learning, effects of competing referents, and error patterns at test) may help to arbitrate among the learning mechanisms proposed by these models.

Another important avenue for future research is to investigate whether the dynamics of the learning process observed here in adults extend to early word learning in children. A number of recent studies have documented that 12- to 18-month-old infants learn cross-situationally in paradigms akin to the ones used in adult word learning studies (e.g., Smith & Yu, 2008; Vouloumanos & Werker, 2009), but it is unclear whether the same constellation of factors reported here also shapes infant word learning. For instance, will infants be more likely to learn words when those words are presented in a more diverse set of contexts? Some evidence (Gomez, 2002; Rost & McMurray, 2009, 2010) suggests that variability facilitates word learning, whereas other evidence suggests that children's word learning is facilitated by limited variability (Maguire, Hirsh-Pasek, Golinkoff, & Brandone, 2008). Thus, extending the current findings to children's cross-situational word learning may shed light not only on the dynamics of early lexical acquisition but also on issues in early learning more broadly.

5. Conclusions

The findings reported here shed light onto the learning processes underlying adult cross-situational word learning. Yu and Smith's clever experimental design lends itself to rich behavioral analyses, which offers a window into these learning dynamics. Our results provide support for Yu and Smith's original claim that a cross-situational statistical learning mechanism, and not a single-trial learning mechanism, underlies performance in this task. Our findings also highlight the complexity of the cross-situational learning process, indicating that multiple sources of information (statistical learning and exclusion constraints) are used conjointly to facilitate learning.

Note

1. When examined at the subject-, rather than the trial-level, the average foil strength for incorrect items ($M = 0.174$, $SD = 0.015$) was marginally larger than correct items ($M = 0.167$, $SD = 0.02$), $t(44) = 1.66$, $p = .10$. These analyses are consistent with the trial-level data. Test foil probability matters, but this is especially true when the highest co-occurring foil is present.

Acknowledgments

References

Akhtar, N., & Montague, L. (1999). Early lexical acquisition: The role of cross-situational learning. *First Language*, *19*, 34–358.

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

Carey, S. (1978). The child as word-learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). Cambridge, MA: MIT Press.

Diesendruck, G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, *37*, 630–641.

Frank, M. C., Goodman, N. D., & Tenenbaum, J. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science*, *20*, 579–585.

Gillette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition*, *73*, 135–176.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*, 3–55.

Golinkoff, R. M., Hirsh-Pasek, K., Bloom, L., Smith, L. B., Woodward, A. L., Akhtar, N., Tomasello, M., & Hollich, G. (2000). *Becoming a word learner: A debate on lexical acquisition*. New York: Oxford University Press.

Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431–436.

Ichinco, D., Frank, M. C., & Saxe, R. (2009). Cross-situational word learning respects mutual exclusivity. In N. Taatgen, H. van Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2214–2219). Austin, TX: Cognitive Science Society.

Kachergis, G., Yu, C., & Shiffrin, R. M. (2009). Frequency and contextual diversity effects in cross-situational word learning. In N. Taatgen, H. van Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2220–2225). Austin, TX: Cognitive Science Society.

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., & Brandone, A. C. (2008). Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science*, *11*, 628–634.

Markman, E. M. (1992). Constraints on word learning: Speculations about their nature, origins, and domain specificity. In M. R. Gunnar & M. Maratsos (Eds.), *Modularity and Constraints in Language and Cognition: The Minnesota Symposia on Child Psychology*, (Vol. 25, pp. 59–101). Hillsdale, NJ: Erlbaum.

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

Pinker, S. (1989). *Learnability and cognition: The acquisition of grammatical structure*. Cambridge, MA: MIT Press.

Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science*, *12*, 339–349.

Rost, G. C., & McMurray, B. (2010). Finding the signal by adding noise: The role of noncontrastive phonetic variability in early word learning. *Infancy*, *15*, 608–636.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science*, *8*, 101–105.

Seger, C. A. (1998) Multiple forms of implicit learning. In M. A. Stadler & P. A. Frensch (Eds.), *Handbook of implicit learning* (pp. 295–320). Thousand Oakes, CA: Sage Publications.

Siskind, J. M. (1996). A computational study of cross-situational techiniques for learning word-to-meaning mappings. *Cognition*, *61*, 39–91.

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

Smith, K., Smith, A. D. M., & Blythe, R. A. (2009). Reconsidering human cross-situational learning capacities: A revision to Yu & Smith's (2007) experimental paradigm. In N. Taatgen, H. van Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2711–2716). Austin, TX: Cognitive Science Society.

Turke-Browne, N. B., Junge, J., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*, 552–564.

Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition*, *107*, 729–742.

Vouloumanos, A., & Werker, J. F. (2009). Infants' learning of novel words in a stochastic environment. *Developmental Psychology*, *45*, 1611–1617.

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

Yu, C., & Ballard, D. H. (2007). A unified model of early word learning: Integrating statistical learning 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. (2011). What you learn is what you see: Using eye movements to study infant cross-situational word learning. *Developmental Science*, *14*, 165–180.

Yu, C., Smith, L. B., Klein, K. A., & Shiffrin, R. M. (2007). Associative learning and hypothesis testing in cross-situational word learning: Are they one and the same? In D. S. McNamara & J. G. Trafton (Eds.) *Proceedings of the 29th Annual Conference of the Cognitive Science Society* (pp. 737–742). Austin, TX: Cognitive Science Society.

Yurovsky, D., & Yu, C. (2008). Mutual exclusivity in cross-situational statistical learning. In B. C. Love, K. McRae, & V. Sloutsky (Eds.) *Proceedings of the 30st Annual Conference of the Cognitive Science Society* (pp. 715–720). Austin, TX: Cognitive Science Society.