

Running head: COMPETITION IN CROSS-SITUATIONAL WORD LEARNING

Competition between multiple words for a referent in cross-situational word learning

Viridiana L. Benitez<sup>a, I</sup>, Daniel Yurovsky<sup>a, II</sup>, and Linda B. Smith<sup>a</sup>

<sup>a</sup>Department of Psychological and Brain Sciences, Indiana University, 1101 E. 10<sup>th</sup> St.,

Bloomington, IN, 47405, USA

Please address correspondence to:

Viridiana L. Benitez

Waisman Center

University of Wisconsin-Madison

Madison, WI, 53705

Email: [vbenitez@wisc.edu](mailto:vbenitez@wisc.edu)

<sup>I</sup> Author's present address: Waisman Center, University of Wisconsin-Madison, Madison, WI, 53705, USA

<sup>II</sup> Author's present address: Department of Psychology, Stanford University, 450 Serra Mall, Stanford, CA, 94305, USA

## Abstract

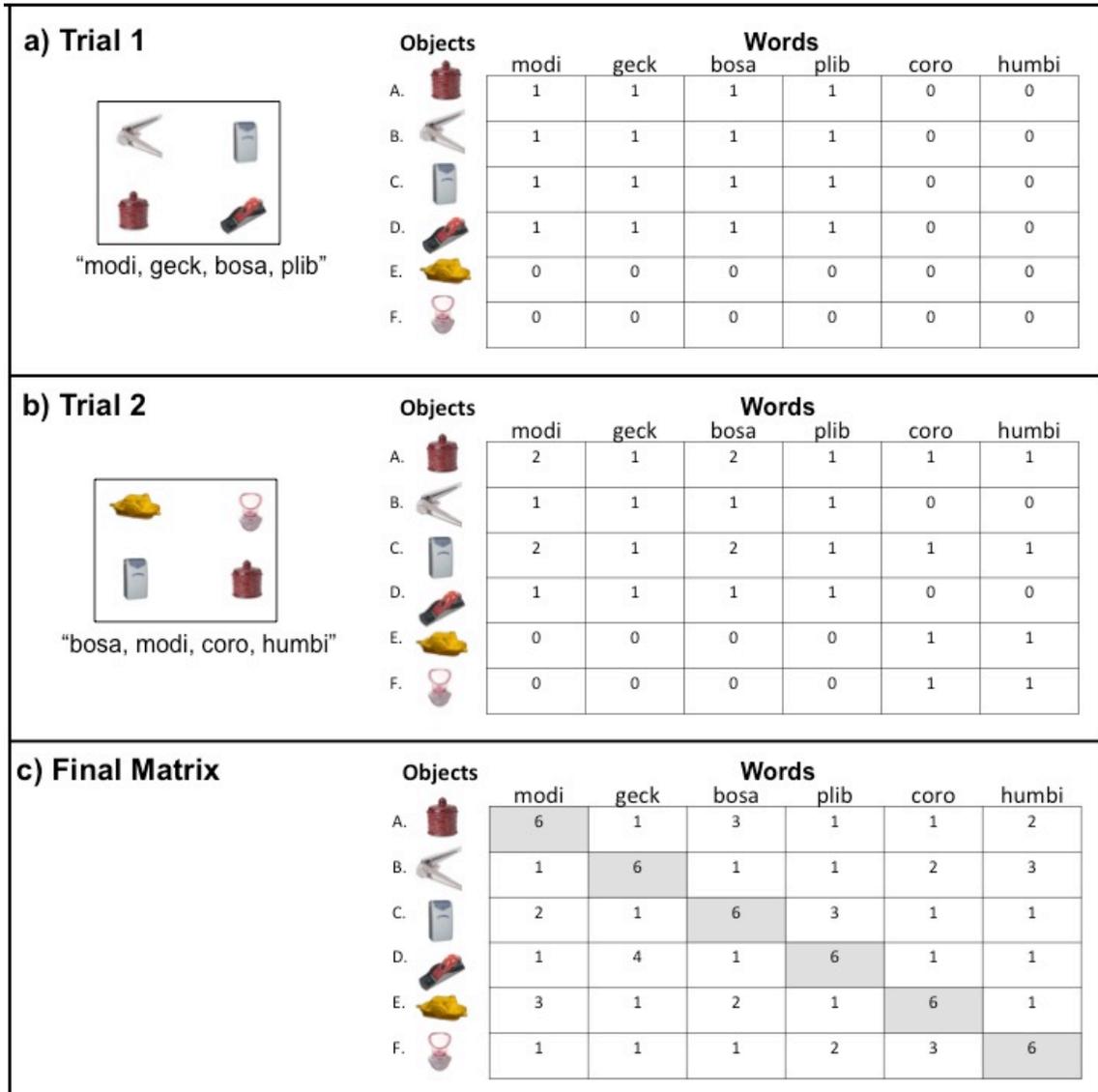
Three experiments investigated competition between word-object pairings in a cross-situational word-learning paradigm. Adults were presented with One-Word pairings, where a single word labeled a single object, and Two-Word pairings, where two words labeled a single object. In addition to measuring learning of these two pairing types, we measured competition between words that refer to the same object. When the word-object co-occurrences were presented intermixed in training (Experiment 1), we found evidence for direct competition between words that label the same referent. Separating the two words for an object in time eliminated any evidence for this competition (Experiment 2). Experiment 3 demonstrated that adding a linguistic cue to the second label for a referent led to different competition effects between adults who self-reported different language learning histories, suggesting both distinctiveness and language learning history affect competition. Finally, in all experiments, competition effects were unrelated to participants' explicit judgments of learning, suggesting that competition reflects the operating characteristics of implicit learning processes. Together, these results demonstrate that the role of competition between overlapping associations in statistical word-referent learning depends on time, the distinctiveness of word-object pairings, and language learning history.

## Competition between multiple words for a referent in cross-situational word learning

Lexical competition is central to many phenomena in language including lexical access and on-line sentence comprehension (e.g., Allopenna, Magnuson, & Tanenhaus, 1998; Cutler, 1995; Levelt, Roelofs, & Meyer, 1999; McClelland & Elman, 1986; Marslen-Wilson, 1990; Norris, 1994). Lexical competition has also been proposed to play an important role in word learning in children and adults (MacWhinney, 1989; McMurray, Horst, & Samuelson, 2012; Merriman, 1999), and is a central mechanism assumed by models of cross-situational word-referent learning (Frank, Goodman, & Tenenbaum, 2009; Kachergis, Yu, & Shiffrin, 2012; Regier, 2005; Siskind, 1996; Smith, Smith, & Blythe, 2011; Yu & Ballard, 2007). Although there is direct evidence of competition in lexical access (Allopenna et al., 1998; Howard, Coltheart, & Cole-Virtue, 2006; Oppenheim, Dell, & Schwartz, 2010), sentence comprehension (Bates & MacWhinney, 1989; Elman, Hare, & McRae, 2005; McRae, Spivey-Knowlton, & Tanenhaus, 1998), and in on-line word-referent disambiguation in children (e.g., Halberda, 2006; Horst, Scott, & Pollard, 2010; Markman, 1990; Merriman, Bowman, & MacWhinney, 1989; Swingley & Aslin, 2007; Yoshida, Tran, Benitez, & Kuwabara, 2011), there is no direct evidence for competition in cross-situational word-referent learning. Here we seek that evidence in the test of one common assumption about how that competition works: individual word-referent associations directly inhibit the pairing of other words with that referent.

The cross-situational word-learning task was designed to measure learners' abilities to find underlying word-referent pairings in the noisy co-occurrence data of heard words and seen things (Yu & Smith, 2007). The task as shown in Figure 1a and 1b

consists of a series of individually ambiguous learning trials in which multiple words and referents are presented with no information about which word goes with which referent. Although individual trials are ambiguous with respect to the word-referent correspondences, each object is always presented with its corresponding word such that, across trials, there is clear evidence for a single set of pairings between words and referents (see Figure 1c). Thus, there is within-trial uncertainty, with many spurious co-occurrences between words and referents, but across-trial consistency, with the strongest co-occurrences indicating the correct word-referent pairings. Studies using this task have shown that adult learners are quite capable –even given many words and referents and after relatively few training trials– of discovering the underlying words and referent pairings from the co-occurrence statistics (e.g., Kachergis et al., 2012; K. Smith, et al., 2011; Suanda & Namy, 2012; Vlach & Sandhofer, 2014; Voloumanos, 2008; Yu & Smith, 2007; Yurovsky, Yu, & Smith, 2013). Even infants and children have been shown capable of learning the word-referent correspondences in these tasks (e.g., Scott & Fisher, 2012; Smith & Yu, 2008; Vlach & Johnson, 2013; Suanda, Mugwanya, & Namy, 2014; Voloumanos & Werker, 2009). To do this, learners must attend to, store, and in some way statistically evaluate the system of word-referent co-occurrences.



**Figure 1.** Example of two trials, their corresponding co-occurrence matrices, and a final matrix, in a typical cross-situational word learning task. a) In Trial 1, four objects are presented with four auditorily presented words with no indication as to which words refer to which objects. At this point, each of the four words have co-occurred with each of the four objects one time, as demonstrated in the co-occurrence matrix. b) In Trial 2, some of the objects presented in Trial 1 are again presented together with objects not yet seen. Here, the words “modi” and “bosa” now have co-occurred with objects A and C twice, and with objects E and F once. Words “coro” and “humbi” have co-occurred once with all objects presented in Trial 2. c) An example of a final co-occurrence matrix after all trials have been presented in a cross-situational word learning task. The co-occurrences reveal the word-referent mappings (the gray cells): “modi” co-occurred most often with object A, “geck” with object B, and so on.

A variety of algorithms, expressed as Bayesian inference models (Frank et al., 2007; Siskind, 1996), machine translation models (Yu & Ballard, 2007), or associative learning models (Kachergis et al., 2012; Regier, 2005), have been shown to be capable of discovering the underlying word-referent pairings from noisy co-occurrence data. A key property of many of these models is that potential word-referent pairings compete.

Within these models, this property of the learning machinery has been shown to be critical to rapid learning (Yu & Smith, 2012) and to the learning of very large sets of words and referents (Blythe, K. Smith, & A.D.M. Smith, 2010; Reisenauer, Smith, & Blythe, 2013; K. Smith et al., 2011). The underlying assumption – implicit in some models, explicit in others (see Yurovsky, et al., 2013) – is that a word-referent pairing with stronger co-occurrence evidence blocks or inhibits the formation of links between other words and that referent. By this assumption, in the matrix of co-occurrences in Figure 1c, earlier co-occurrence data between the word “modi” and object A should inhibit the later pairing of another name, e.g., “bosa”, to object A, with the resolution of this competition being a function of the relative associative strength of the two competing items. This proposed competition component in cross-situational learning is similar to competition processes found in several prominent models of word learning more generally (MacWhinney, 1989; McClelland & Elman, 1986; McMurray et al., 2012).

Such competitive processes make a strong prediction: there should be direct competition at the item level between specific words that share a referent. For example, if learners acquire one word-referent pairing strongly, learning another word for that referent should be more difficult. To date, although a variety of models that propose item-level competition have been fit to learning data, item-level competition itself has not been

empirically demonstrated. The main goal of the following three experiments was to document item-level competition; a secondary goal was to determine possible limits on inter-item competition with the aim of providing potential insights as to the mechanisms or stages of learning at which competition occurs.

To these ends, we used a variant of the cross-situational word learning task shown in Figure 1, but in our version, shown in Table 1, some referents were principally associated with one word and other referents were equally associated with two words (see also Ichinco, Frank, & Saxe, 2009; Kachergis et al., 2012; Poepsel & Weiss, 2014; Yurovsky et al., 2013). More specifically, for the One-Word pairings, a single word co-occurred every time with its object, and the frequency and probability of these co-occurrences were much greater than the spurious co-occurrences of that object with other words (e.g., word d with object D, see Table 1). For Two-Word pairings, each object (e.g., object A) co-occurred equally and most often with two words (e.g., word a1 and word a2). Previous research has shown that these statistics should result in weaker learning of Two-Word pairings relative to One-Word pairings (Ichinco et al., 2009; Kachergis et al., 2012; Yurovsky et al., 2013) a fact that might seem to suggest direct inter-item competition. However, the advantage of input statistics that favor One-Word pairings over Two-Word pairings may also be explained by other processes, such as differences in conditional probabilities. While some studies have attempted to control for this by manipulating the conditional probabilities (e.g., Kachergis et al., 2012; Yurovsky et al., 2013), none have directly assessed the degree of competition among individual overlapping items. Thus, in addition to assessing overall performance on overlapping pairings as others have done (Ichinco et al., 2009; Kachergis et al., 2012; Yurovsky et al.,

2013), we assessed individual trial data at testing to obtain a direct measure of competition between overlapping pairings.

**Table 1**  
Co-occurrence matrix used in Experiments 1, 2, and 3

Objects	Words																	
	a1	b1	c1	d	e	f1	g1	h1	i	j	k	l	g2	c2	b2	a2	f2	h2
A	6	3	3	2	2	1	2	1	3	2	3	3	1	4	3	6	2	1
B	3	6	2	1	2	3	0	1	4	1	2	1	4	4	6	3	2	3
C	4	4	6	1	1	2	2	2	1	3	4	1	3	6	2	3	1	2
D	0	0	1	6	3	1	1	2	0	1	0	2	2	0	1	2	0	2
E	1	0	1	3	6	0	1	2	1	1	1	1	0	0	2	1	2	1
F	1	3	2	1	2	6	2	1	3	4	3	4	2	1	2	2	6	3
G	2	2	2	3	1	2	6	4	2	2	2	3	6	3	2	1	2	3
H	0	2	2	4	3	3	4	6	2	2	2	2	3	2	2	2	1	6
I	3	3	0	0	1	1	1	1	6	2	0	0	1	1	1	0	2	1
J	1	1	2	1	1	1	2	2	2	6	0	0	0	1	0	1	3	0
K	2	0	3	0	1	1	1	1	0	0	6	1	1	1	2	1	2	1
L	1	0	0	2	1	3	2	1	0	0	1	6	1	1	1	2	1	1

According to the principle of relative-strength competition (Desimone & Duncan, 1995; Mesink & Raijmakers, 1988; Norman, Newman, & Detre, 2007), the degree and the resolution of competition should be a function of the relative strength of the competing items. To the extent that one pair is well learned, its overlapping competitor should be poorly learned. This is the prediction of item-level competition that is tested in the following experiments: If items directly compete, the learning of one word for an object should be negatively related to the learning of another word for that same object. Furthermore, if this competition is based on the strength of competing items, then time and the distinctiveness of individual pairings may affect the presence of competition

(e.g., Estes, 1989). We tested the effect of time on competition by presenting the co-occurrences of overlapping pairings interleaved (Experiment 1), or blocked (Experiment 2) during training. In Experiment 3, we tested the role of association distinctiveness for competition by adding a linguistic cue to overlapping pairings.

Recent research (see, Perruchet & Pacton, 2006, for a review) on cross-situational word-referent learning, and statistical learning more generally, has asked whether the learning mechanisms in these tasks are a form of implicit learning (Reber, 1993), a question relevant to the contrasting views of word-referent learning as hypothesis testing or a form of associative learning (see Yu & Smith, 2012, for discussion). Accordingly, we asked participants on each training trial to indicate their confidence that they knew which word went with which object on that trial, a measure of their awareness of how well they were learning. A participant's confidence on a given training trial could be based on spurious co-occurrences, and, if learning is principally implicit, confidence ratings need not indicate (or even correlate with) participants' knowledge of correct co-occurrences. The same indirect approach procedure was used by Yurovsky et al. (2013, Experiment 4; see also Poepsel & Weiss, 2014), to assess participant's trial-by-trial approach to the task without asking them to explicitly link words and objects, since explicit responses may in and of themselves influence the registration of word-object correspondences (Fitneva & Christiansen, 2011, in press; Turk-Browne, Scholl, Johnson, & Chun, 2010).

Finally, in all experiments, we included an individual-difference measure of self-reported language experiences, dividing participants into two groups –those with experience learning multiple languages and those with experiences learning only one

language (English). We did this for two reasons: First, the key manipulation, two words versus one word associated with the same object, could be viewed as creating training sets that are like learning multiple languages. Second, there have been mixed reports in the literature that experiences with multiple languages may alter proficiency in cross-situational word learning tasks (Escudero, Mulak, & Vlach, 2015; Poepsel & Weiss, submitted). However, in other task contexts, the evidence has been more consistent: experience with multiple languages has been shown to affect word learning (Brojde, Ahmed, and Colunga, 2012; Byer-Heinlein & Werker, 2009; Kaushanskaya & Marian, 2009; Yoshida et al., 2011), statistical learning (Bartolotti, Marian, Schroeder, & Shook, 2012; Wang & Saffran, 2014), and the resolution of competition during high conflict tasks (Adesope, Lavin, Thompson, & Unglreider, 2010; Hilchey & Klein, 2011; Kroll & Bialystok, 2013; Valian, 2015). Thus, previous language learning experience is a factor that could affect how competition between word-referent pairings is resolved, a finding that would implicate individual history-malleable effects on competitive processes. Alternatively, fundamental effects of competition may be more firmly fixed in the learning machinery. As a first step in answering these larger questions, we assessed language-learning history through a self-report measure of language-learning experiences: (1) by comparing participants who reported learning (and speaking) multiple languages to those who only reported speaking one, and (2) by examining the extent of second language experience for those who reported experiences learning multiple languages.

## Experiment 1

Experiment 1 presented learners with One-Word and Two-Word pairings intermixed in training. Thus, the input evidence for overlapping word-referent pairings could increase similarly across trials. Although the input evidence for one member of an overlapping pair was no greater than the evidence for the other, if the items compete, then knowledge of the two items should be negatively related. Small early advantages for one member in the random mix of training trials should disadvantage learning the other and the resolution of competition more in favor of one versus the other competitor should increase across trials.

### Method

**Participants.** Participants for this experiment were recruited through the use of both the student experiment system and flyers posted throughout the university campus to increase the range of participants. The participants were 103 adults who ranged in education level from first year college students to a graduate degree. Participants received course credit or pay (\$10) for participating in the experiment. We collected information from all participants about their language experiences using the Language History Questionnaire (Li, Sepanski, & Zhao, 2006). Using the answers from this questionnaire, we partitioned subjects into broad groups. The One Language group (N=50, 19 male) was defined as individuals who indicated they did not speak a second language and thus had minimal experiences learning other languages. The Multiple Languages group (N=43, 14 male) was defined as anyone who indicated they spoke more than one language and thus had more extensive experiences in learning other languages. The One Language participants spoke only English, the Multiple Languages participants spoke English and

at least to some degree some other language. Table 2 provides the demographic and linguistic details of the participants. There were no significant differences between groups on any of the demographic factors other than country of birth: about half of the Multiple Languages group were foreign born, while none were for the One Language group.

**Table 2**  
Demographic and linguistic details of One Language and Multiple Languages groups in Experiments 1-3

	Experiment 1		Experiment 2		Experiment 3	
	One <sup>a</sup>	Multiple	One	Multiple	One	Multiple
<b>Age</b>						
Mean (SD)	21.34(8.1)	21.43(3.77)	19.52(2.06)	21.54(3.83) <sup>b</sup>	23.15(10.3)	26.6(12.1)
Range	18-57	18-35	18-27	18-34	18-64	18-64
<b>Education</b>						
Undergraduate or Some College	0.94	0.74	0.90	0.76	0.96	0.98
Bachelor's	0.04	0.15	0.10	0.12	0	0
Masters or higher	0.02	0.11	0	0.12	0.04	0.02
<b>Country of Origin</b>						
Foreign Born	0	0.53	0	0.6	0	0.5
<b>Number of Languages Known</b>						
Mean(SD)	-	2.15(0.63)	-	2.14(0.35)	-	2.29(0.77)
Range	-	2-6	-	2-3	-	2-6
Proportion with more than 2	-	0.08	-	0.14	-	0.17
<b>Age of exposure to 2nd language</b>						
Mean(SD)	-	6.27(5.17)	-	6.58(4.41)	-	6.98(5.4)
Range	-	0-21	-	0-16	-	0-19
Proportion before age 10	-	0.74	-	0.7	-	0.63
<b>Proficiency in 2nd language (scale of 1-7, 7 being native-like)</b>						
Mean(SD)	-	6.13(1.04)	-	5.93(1.38)	-	6.0(1.12)
Range	-	3-7	-	1-7	-	3-7
Proportion 6 or higher	-	0.75	-	0.76	-	0.73

<sup>a</sup>Four participants in the One Language group did not report age and education

<sup>b</sup>The Multiple Languages group was statistically older than the One Language group

**Stimuli.** The stimuli consisted of two sets of 18 unfamiliar objects and 18 novel words that followed English phonotactic rules. These were generated in a synthetic, monotone, female voice using the AT&T Natural Voices® system. Figure 2 depicts one set of stimuli. Participants were randomly assigned to one set, out of which twelve objects and 18 words were chosen and randomly paired (for each participant) to yield 6 One-Word pairings, generated by pairing a unique single word with a unique single object, and 6 Two-Word pairings, generated by pairing two unique words with a single unique object. Each Two-Word pairing could be broken down into two separate word-object pairings, Word 1 and Word 2, each composed of the same object labeled with a different word (e.g., a1—A had the corresponding pairing a2—A). This design is similar to Yurovsky et al. (2013), except here, objects were labeled with two words, whereas in Yurovsky et al. (2013), a single word labeled two objects.




---

Experiment 1 & 2

<i>gasser</i>	<i>coro</i>	<i>flib</i>	<i>bosa</i>	<i>habble</i>	<i>humbi</i>	<i>kaki</i>	<i>regli</i>	<i>permi</i>
<i>gaso</i>	<i>toma</i>	<i>ceturg</i>	<i>temi</i>	<i>minick</i>	<i>coston</i>	<i>fena</i>	<i>geck</i>	<i>blandi</i>

---

Experiment 3

Disyllabic with vowel ending	<i>baka</i>	<i>poda</i>	<i>gaso</i>	<i>mapu</i>	<i>kumbi</i>	<i>detee</i>	<i>scabby</i>	<i>temi</i>	<i>brenoe</i>
	<i>brarry</i>	<i>tera</i>	<i>toma</i>						
Monosyllabic with /k/ ending	<i>meek</i>	<i>cusk</i>	<i>sprock</i>	<i>glark</i>	<i>jick</i>	<i>roak</i>			

**Figure 2.** One set of 18 unfamiliar objects and two sets of 18 novel words used in the cross-situational word-learning tasks in Experiment 1, 2, and 3. All novel words followed English phonotactic rules, and were generated using a synthetic female voice. Participants were randomly assigned to one of two sets. In Experiments 1 and 2, words were randomly paired for each participant to generate 6 One-Word pairings (where a single label consistently co-occurred with a single object) and 6 Two-Word pairings (where two words consistently co-occurred with the same object). In Experiment 3, all words for One-Word objects were disyllabic with a vowel ending. Two-Word objects had one word that was of this structure, and a second word that was monosyllabic with the stop consonant ending /k/.

**Training.** Training trials were set up so that each trial presented participants with four objects (placed on the four corners of the computer screen) and four words. The objects were presented on the screen first, and approximately 2 seconds after, the four

words were auditorily presented, each separated by approximately 2 seconds. After all four objects and all four words were presented for an individual trial, participants were asked to rate, for each object, on a scale of 1 to 10, how confident they were that they knew the name for that object (1 being not at all confident, 10 being absolutely confident).

In training, each word appeared consistently with its object 6 times, including each word in Two-Word pairings (e.g., word a1 appeared 6 times with object A, and word a2 appeared 6 times with object A). Thus, adults received six exposures of each of six One-Word pairings, and six exposures for each word for six Two-Word objects. These pairings were presented in a randomly intermixed fashion in a total of 27 training trials. Individual training trials could include all One-Word pairings, individual Two-Word pairings, or some combination of these two. The two words for Two-Word objects were never presented on the same trial, and instead, they were randomly interleaved between trials in training (e.g., a1—A could be presented on trial 1, then a2—A on trial 2, and again on trial 4, and then a1—A again on trial 7). This design was implemented to avoid attentional competition within individual training trials. There were on average 1.77 repetitions of an individual word before the other word for that same object was presented.

The co-occurrence statistics at the end of training are shown in Table 1. Notice that the frequency of co-occurrence for correct word-object pairings for One-Word and for both words of a Two-Word object are the same. However, the conditional probabilities differ across One-Word and Two-Word pairings. For One-Word pairings, when an object was presented, the probability of hearing its respective word was 1.0.

However, for Two-Word pairings, when an object was seen, each of its respective words was heard with a probability of 0.5. If subjects compute conditional probabilities, then one would expect lower performance on the Two-Word pairings relative to the One-Word pairings, regardless of item-level competition. Thus overall performance on the One-Word versus Two-Word pairings does not provide a test of item-level competition. Rather, the test for item-level competition consists of comparisons between words that share the same referent (overlapping Two-Word pairings), which have equivalent conditional probabilities.

**Testing.** For test, we used a four alternative forced choice procedure. Test trials consisted of the auditory presentation of one word and the visual presentation of four objects on the screen. Participants were presented with a total of 18 test trials. To limit effects of the experiences during testing on performance on individual items, there was one test trial for each word. Objects that were paired with two words were thus the correct choice on two test trials, once for each of their words. The three distractor objects for each test trial were randomly chosen. Test trial order was also randomized.

**Procedure.** To complete the training portion of the task, participants were instructed to learn which words refer to which objects without being told anything about the structure of the co-occurrences. Testing immediately followed training, and to complete the testing trials, adults were instructed that when they heard a word, to choose the object to which they thought the word referred. The language questionnaire was filled out after the cross-situational word-learning task. The total session lasted approximately 30 minutes.

## **Results and discussion**

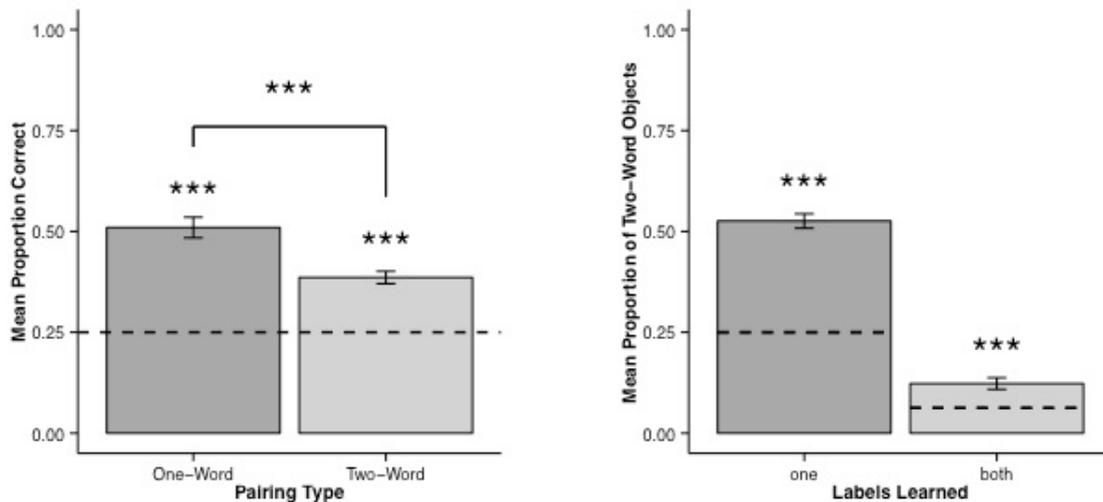
We first calculated the overall proportion of correct responses for One-Word and Two-Word pairings. For Two-Word pairings, we aggregated across all 12 words to obtain a measure of overall accuracy for Two-Word pairings (without taking into account corresponding pairings).

Figure 3a shows what prior experiments have shown (see Ichinco et al., 2009; Kachergis et al., 2012; Yurovsky et al., 2013): participants were above chance on both pairings types, but they learned the One-Word pairings better than the Two-Word pairings. A 2 (Pairing Type) by 2 (Language Group) mixed measures ANOVA showed that One-Word pairings were learned significantly better than Two-Word pairings [by-subjects:  $F(1, 101) = 27.14, p < 0.001, 95\% \text{ CI } [0.09, 0.34], \eta^2_G = 0.079$ ; a by-item ANOVA revealed the same pattern]. Participants had equal numbers of encounters with all of the individual associations – in both One-Word and Two-Word pairings – however, there were clear differences in learning between the two pairing types. This result is consistent with past research showing better learning of pairings that have less overlap with other pairings in their co-occurrence statistics than pairings with higher overlap (see Ichinco et al., 2009; Kachergis et al., 2012; Yurovsky et al., 2013).

The advantage of One-Word pairings over Two-Word pairings was also consistent across the two language groups, as there was no effect of Language Group [by-subject:  $F(1, 101) = 0.12, p = 0.73$ ], nor an interaction [by-subject:  $F(1, 101) = 0.00034, p = 0.99$ ; the by-item ANOVA revealed the same result]. Participants in both groups performed above chance in both the One-Word and Two-Word conditions. [One Language group: One-Word pairings,  $M = 0.50, SD = 0.28, t(49) = 6.43, p < 0.001, 95\% \text{ CI } [0.42, 0.58], d = 0.91$ ; Two-Word pairings:  $M = 0.38, SD = 0.15, t(49) = 5.99,$

$p < 0.001$ , 95% CI [0.34, 0.42],  $d = 0.85$ ; Multiple Languages group: One-Word Pairings,  $M = 0.52$ ,  $SD = 0.24$ ,  $t(52) = 8.03$ ,  $p < 0.001$ , 95% CI [0.45, 0.58],  $d = 1.10$ ; Two-Word Pairings,  $M = 0.39$ ,  $SD = 0.16$ ,  $t(52) = 6.55$ ,  $p < 0.001$ , 95% CI [0.35, 0.43],  $d = 0.90$ ].

Additionally, differences in the self-reported language learning experience in the Multiple Languages group also did not have an effect on overall learning of One-Word and Two-Word pairings as indicated by 3 different measures: self-rated proficiency (self-rated Native-like,  $N=25$ , versus less than Native-like,  $N=24$ ); age of first exposure (before 5 years,  $N=30$ , versus 5 years or older,  $N=23$ ); English ( $N=29$ ) as a first or second language ( $N=24$ ). Separate 2 (Pairing type) by 2 (Group) ANOVAs for each measure indicated no reliable differences. Accordingly, language history was not included in further analyses of Experiment 1.



**Figure 3.** a) Means, and standard errors of the mean for accuracy for objects labeled with a single word (One-Word pairings) and objects labeled with two words (Two-Word pairings) in a cross-situational word-learning task where these pairings were presented intermixed during training (Experiment 1). b) Proportion of Two-Word objects for which one label or both labels were learned in Experiment 1. For both graphs, dashed lines denote chance performance and asterisks denote significant differences at the  $p < 0.001$  level. Note that performance here is collapsed across Language group, as there was no difference between One Language and Multiple Languages learners in Experiment 1.

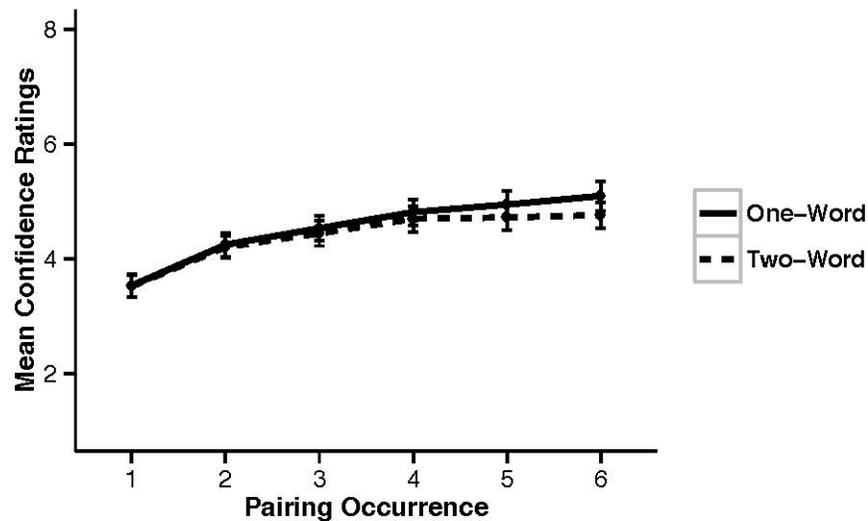
The main goal of this experiment was to measure inter-item competition. Performance on the individual competing pairings within the Two-Word pairings (e.g. if learning a1—A inhibited learning a2—A) provides the critical information about competition. There were a total of 12 Two-Word tests, one for each word for the 6 Two-Word objects. Figure 3b displays the proportion of Two-Word objects for which just one label was learned as compared to the proportion of Two-Word objects for which both labels were learned. Participants learned one label and both labels at above chance levels [one label:  $t(102) = 15.52$ ,  $p < 0.001$ , 95% CI [0.49, 0.56],  $d = 1.53$ ; both labels:  $t(102) = 4.21$ ,  $p < 0.001$ , 95% CI [0.09, 0.15],  $d = 0.42$ ]. They also were more likely to learn one label for an object than both labels [ $t(102) = 6.12$ ,  $p < 0.001$ , 95% CI [0.35, 0.45],  $d = 2.46$ ].

If competition exists between words that label the same object, then the strength of knowledge of one item should be *inversely* related to its overlapping competitor. Accordingly, we took accuracy scores (either 1 or 0) for individual Two-Word object test trials, and we used a mixed effects logistic regression (with random effect terms of subject and item) to estimate the predictability of Word 1 accuracy for Word 2 accuracy. Note that the two labels for each Two-Word object were randomly assigned as Word 1 or Word 2. The logistic regression coefficient was negative and significant, showing that if participants were accurate on the test for Word 1, the likelihood of them being accurate on the test for Word 2 decreased [Beta = -0.51, odds ratio = 0.60, STE = 0.23,  $p = 0.03$ , 95% CI [-0.95, -0.06]]. This strong link between Word 1 and Word 2 was also present when we defined Word 1 as the first word presented at training, or the first word presented at test. This result provides direct support for the hypothesis that knowledge of

one word for a Two-Word object competed with knowledge of the other word for that object. This is the first evidence of item-level competition in the learning of co-occurrence statistics and it appears to be a general one, across a wide range of participants: Words that refer to the same object compete such that knowledge of one is negatively related to knowledge of the other.

How does this competition relate to participants' awareness of their learning across trials? One possibility is that a noticed co-occurrence sufficiently inhibits competing words for the same object such that participants do not even realize that there is competition, and thus are equally confident during training on Two-Word trials as on One-Word trials. Alternatively, the co-occurrence data for the Two-Word items-- that two words map on to one object -- may be explicitly noticed and lead to less overall confidence on Two-Word than One-Word items. The results of participants' confidence ratings during learning, shown in Figure 4, are more in line with the first hypothesis. A repeated measures ANOVA (by-subject), with two within-subject factors of Occurrence and Pairing Type revealed only a main effect of Occurrence [ $F(5, 510) = 37.73, p < 0.001, 95\% \text{ CI } [0.20, 0.32], \eta^2_G = 0.044$ ], where participants' confidence increased throughout training. The effect of Pairing Type [ $F(1, 102) = 3.02, p = 0.085$ ] was not reliable, and did not interact with Occurrence [ $F(5, 510) = 1.57, p = 0.166$ ; a by-item ANOVA revealed the same set of patterns]. Relative to their performance at test, participants were over-confident of their knowledge of Two-Word pairings. To assess this over-confidence, we used a logistic regression (with random effect terms of subject and item) to ask how confidence ratings on each occurrence of a pairing predicted performance at test for that pairing. Table 3 presents the regression coefficients for each occurrence for each pairing

type, showing that by the last occurrence of word-object pairings, only One-Word pairings were significantly predictive of testing performance. Overall the results and the confidence ratings are consistent with the idea of active competition during learning itself. Learners were selectively forming one-word-one-object mappings, inhibiting the registration of competing co-occurrences, and thus only learning one word for a Two-Word object while unaware of a second label.



**Figure 4.** Mean ratings (and standard errors of the mean) for One- and Two-Word pairings for each pairing occurrence in training for Experiment 1 collapsed across Language group. Participants were asked to rate, from 1-10, their knowledge of the word for each object (1 meaning I do not know the name for this object, and 10 meaning I do know the name for this object). There was no statistical difference in confidence ratings for One-Word and Two-Word pairings.

**Table 3**

Regression coefficients (Betas) for logistic regressions predicting test accuracy from confidence ratings on each occurrence (1-6) of a word-object pairing

Pairing Type	1	2	3	4	5	6
<b>Experiment 1</b>						
One-Word	-0.053	-0.037	-0.021	0.014	0.081	0.16***
Two-Word	-0.074*	0.024	0.022	-0.014	0.074*	0.043
<b>Experiment 2</b>						
One-Word	0.024	-0.025	0.071	0.031	0.023	0.22***
Two-Word 1st	-0.10*	-0.026	-0.010	0.12*	0.099	0.14**
Two-Word 2nd	-0.12*	0.043	-0.028	0.057	0.097*	0.05
<b>Experiment 3</b>						
<b>One-Language</b>						
One-Word	0.026	-0.0091	0.069	0.10	0.0082	0.042
Two-Word Cued	-0.0016	-0.052	0.016	0.10	0.029	0.036
Two-Word Uncued	-0.14	0.13	-0.029	0.047	-0.074	0.084
<b>Multiple-Languages</b>						
One-Word	0.055	0.025	-0.040	0.048	0.019	0.10
Two-Word Cued	0.0072	0.036	-0.029	-0.011	0.071	0.11†
Two-Word Uncued	-0.088	0.077	-0.091	0.099	-0.013	0.23**

Note: In Experiments 1 and 2, there was no difference in overall performance across the One and Multiple Languages groups, so only the overall data patterns are presented.

† $p < 0.07$ . \* $p < 0.05$ . \*\* $p < 0.01$ . \*\*\* $p < 0.001$

In sum, the results of Experiment 1 support the hypothesis of item-level competition such that evidence for one word-object association interferes with learning about a competing word-object association. Further, this competition appears unrelated to participants' trial-by-trial awareness of the multiple labels for some objects and unrelated to self-reported language learning history.

### Experiment 2

Experiment 2 separated participants' experiences of the overlapping pairings into two distinct training blocks. Participants were presented with one word for each referent in a first training block followed by a second training block with a new word for some of

those referents. There are two opposing predictions about how blocked presentations should influence inter-item competition. The first prediction derives from the idea that a strongly learned word for a referent will inhibit the learning of a new word for that referent. This kind of competition is reflected in models and evidence of mutual exclusivity (Ichinco, Frank, & Saxe, 2009; Yurovsky et al., 2013), indicating a bias to learn one-to-one word referent pairings. If this form of competition underlies learning, then blocking each word for a referent should result in increased competition. The first-presented pairings should inhibit learning of the second-presented pairings. The second and opposite prediction is that blocking the two words for an object eliminates competition. This prediction derives from the literature on interleaved versus blocked training trials. In category and memory paradigms, interleaving instances over time has been shown to benefit learning relative to massing (or blocking) instances (e.g., Cepeda, Pashler, Vul, & Rohrer, 2006; Ebbinghaus, 1913). However, the evidence in cross-situational word learning suggests that the opposite may be the case, and that blocking may be more beneficial (Vlach & Johnson, 2013; Poepsel & Weiss, 2014; Yurovsky et al., 2013). Blocking word-object pairings may benefit statistical learning because proximity in time supports aggregation across repetitions of the very same word-referent pairing (Vlach & Johnson, 2013). In the same way, blocking may reduce competition between overlapping pairings by allowing their aggregation within a block without interference from the overlapping pairing.

## **Method**

**Participants.** Participants were 95 adults who did not participate in Experiment 1, and the same self-report measures were used as in Experiment 1 to group them into the

One Language group (N=46, male = 12) and the Multiple Languages group (N=50, male =21). Demographic and linguistic details of the participants are presented in Table 2.

Participants in the Multiple Languages group were slightly (though statistically significant) older than the One Language group [ $t(94) = -3.18, p = 0.002$ ].

**Stimuli.** The stimuli consisted of the same 12 unfamiliar object pictures and 18 pseudowords in Experiment 1 (see Figure 2). Just as in Experiment 1, participants were randomly assigned to one of two sets, and the objects and words in each set were randomly matched up for each participant to yield 6 One-Word pairings and 6 Two-Word pairings.

**Training.** Training trials were set up similarly to Experiment 1 with the only difference being the blocked presentations of words that labeled the same object. During Block 1, co-occurrences between labels and objects followed the structure of one-to-one mappings: all 12 objects were presented, and they co-occurred most often with a *single* word. In Block 2 of training, half of those objects continued to be labeled each with their single word presented in Block 1 (One-Word pairings), and each of the other half of the objects were labeled with a novel, second word (and no longer labeled with the first word presented in Block 1; Two-Word pairings). Thus, all 6 co-occurrences of Word 1 for a Two-Word object were presented in Block 1 (Two-Word 1<sup>st</sup>), all 6 co-occurrences of Word 2 were presented in Block 2 (Two-Word 2<sup>nd</sup>), and the 6 co-occurrences for each One-Word pairing were distributed across the two blocks (1-3 in Block 1, 4-6 in Block 2). The presentation of each individual training trial, including timing, confidence ratings, and number of training trials, was the same as Experiment 1. It is also important to note that once more this design included the differences in conditional probabilities between

One- and Two-Word pairings that existed in Experiment 1. The co-occurrence matrix for Experiment 2 was the same as Experiment 1 (see Table 1).

**Testing.** Testing trials were set up exactly the same as Experiment 1.

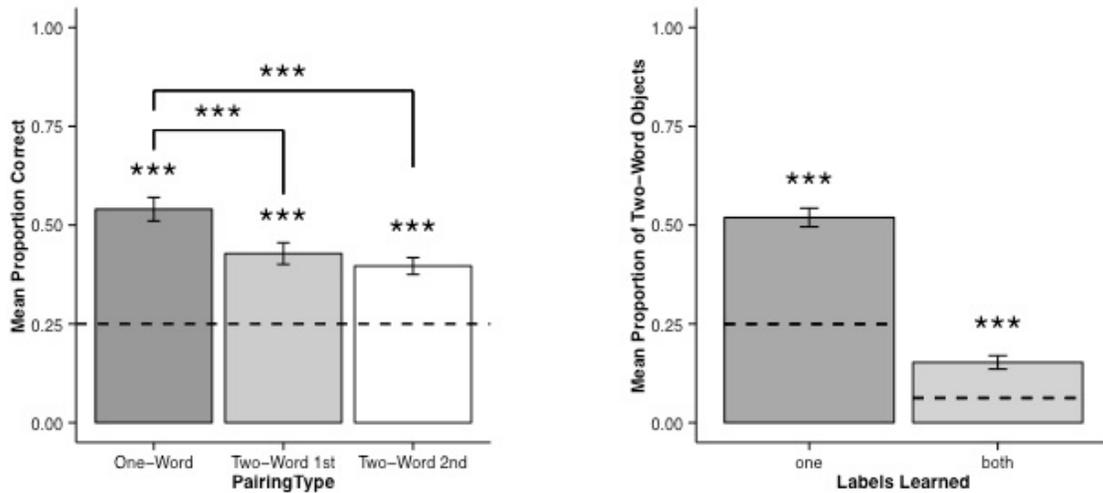
**Procedure.** The instructions and procedure were the same as Experiment 1. The total session lasted approximately 30 minutes.

### Results and discussion

We calculated the overall proportion of correct responses for each pairing type separately (One-Word pairings, Two-Word 1<sup>st</sup> pairings, and Two-Word 2<sup>nd</sup> pairings). A mixed measures ANOVA (by-subject) showed only a significant effect of Pairing Type ( $F(2, 186) = 11.11, p < 0.001, 95\% \text{ CI } [0.03, 0.19], \eta^2_G = 0.055$ ), with no effect of Language Group [ $F(1, 93) = 0.35, p = 0.56$ ] or an interaction [ $F(2, 186) = 0.48, p = 0.62$ ; a by-item ANOVA also showed the same results]. As shown in Figure 5a, One-Word pairings were learned better than Two-Word pairings, and there was no difference between Two-Word 1<sup>st</sup> and Two-Word 2<sup>nd</sup> pairings. Independent samples t-tests (with Bonferroni corrections) demonstrated that overall, One-Word pairings were learned significantly better than Two-Word 1<sup>st</sup> pairings [ $t(94) = 3.53, p = 0.002, 98.3\% \text{ CI } [0.03, 0.19], d = 0.40$ ] and Two-Word 2<sup>nd</sup> pairings [ $t(94) = 4.72, p < 0.001, 98.3\% \text{ CI } [0.07, 0.22], d = 0.57$ ]. The Two-Word 1<sup>st</sup> and 2<sup>nd</sup> pairings, however, were not significantly different from each other [ $t(94) = 0.94, p = 1.0$ ]. Recall that the final co-occurrence frequency (see Table 1) is the same for each individual One-Word pairing and each individual pairing of the Two-Word set, and that for the first half of training, Two-Word 1<sup>st</sup> items were more frequent than the One-Word items. The finding that participants learned the One-Word items better than the Two-Word 1<sup>st</sup> pairings clearly shows an

effect of the lack of continuing evidence for Two-Word 1<sup>st</sup> items; the fact that One-Word items were better learned than Two-Word 2<sup>nd</sup> items shows that the learning of these items was affected by prior experience. A lack of a strong order effect with no advantage of the Two-Word 1<sup>st</sup> items over the Two-Word 2<sup>nd</sup> items shows that strong initial evidence for one pairing does not limit learning of the second, when these are learned in separate training blocks.

Performance was above chance for all three pairing types for both language groups [One Language group: One-Word,  $M = 0.51$ ,  $SD = 0.30$ ,  $t(44) = 5.83$ ,  $p < 0.001$ , 95% CI [0.42, 0.60],  $d = 0.87$ ; Two-Word 1<sup>st</sup>,  $M = 0.42$ ,  $SD = 0.28$ ,  $t(44) = 4.09$ ,  $p < 0.001$ , 95% CI [0.34, 0.50],  $d = 0.61$ ; Two-Word 2<sup>nd</sup>,  $M = 0.40$ ,  $SD = 0.20$ ,  $t(44) = 5.14$ ,  $p < 0.001$ , 95% CI [0.34, 0.46],  $d = 0.77$ ; Multiple Languages group: One-Word,  $M = 0.57$ ,  $SD = 0.28$ ,  $t(49) = 7.89$ ,  $p < 0.001$ , 95% CI [0.49, 0.65],  $d = 1.12$ ; Two-Word 1<sup>st</sup>,  $M = 0.44$ ,  $SD = 0.26$ ,  $t(49) = 5.12$ ,  $p < 0.001$ , 95% CI [0.36, 0.51],  $d = 0.72$ ; Two-Word 2<sup>nd</sup>,  $M = 0.39$ ,  $SD = 0.22$ ,  $t(49) = 4.66$ ,  $p < 0.001$ , 95% CI [0.33, 0.46],  $d = 0.66$ ]. As in Experiment 1, the lack of self-reported language history effects was confirmed by examining differences as a function of three additional measures in the Multiple Languages group (defined as in Experiment 1: Native-like proficiency in the second language, Age of Exposure, and English as a native language). Again, there were no reliable differences in performance on any of the pairings for these Multiple Languages subgroups, so the factor of language was excluded from any further analyses in Experiment 2.



**Figure 5.** a) Means and standard errors of the mean for accuracy for One- and Two-Word pairings for Experiment 2, where the first and second labels of Two-Word objects were blocked in training. b) Proportion of Two-Word objects for which one or both Labels were learned in Experiment 2. Here, as in Experiment 1, data are collapsed across the One Language and Multiple Languages groups, as there were no differences between the two. For both graphs, dashed lines denote chance performance and asterisks denote significant differences at the  $p < 0.001$  level.

Do overlapping pairings compete across the two blocks? The fact that Two-Word 1<sup>st</sup> and 2<sup>nd</sup> pairings are learned equally well does not in-and-of itself rule out competition. Both of these types of pairings were still learned less well than One-Word pairings, and this difference could have been due to direct competition. On the second half of the trials, participants may have retained some Two-Word 1<sup>st</sup> pairings and inhibited some Two-Word 2<sup>nd</sup> pairings, or, as they registered new Two-Word 2<sup>nd</sup> pairings, they may have inhibited previously registered pairings. These processes may not be mutually exclusive, which may have led to both types of Two-Word pairings being learned less well than One-Word pairings. The key prediction is this: if overall less accurate learning in Two-Word pairings (as compared to One-Word pairings) is due to item-level competition, then

knowledge of corresponding Two-Word 1<sup>st</sup> and Two-Word 2<sup>nd</sup> pairings should be negatively related.

Alternatively, the overall learning pattern in Figure 5a could emerge without direct competition. Instead, the advantage of One-Word versus Two-Word pairings could reflect the over-the-whole experiment differences in conditional probabilities linking words to objects. By design, conditional probabilities for One-Word pairings were higher than conditional probabilities for Two-Word pairings. Adults may have been tracking these across the two blocks, leading to better performance for One-Word items. Better learning of One-Word pairings than Two-Word pairings could also emerge without item competition if participants were only learning *some* of the One-Word and Two-Word 1<sup>st</sup> pairings during the first half of the experiment, and then adding more items of One-Word pairings, and *some* items of Two-Word 2<sup>nd</sup> pairings. This pattern would lead to –in the end –learning more One-Word items than either the Two-Word 1<sup>st</sup> and Two-Word 2<sup>nd</sup> items. If these two possibilities are driving the differences between One-Word and Two-Word pairings, without item-level competition, we would expect no relation between the learning of the corresponding Two-Word pairings.

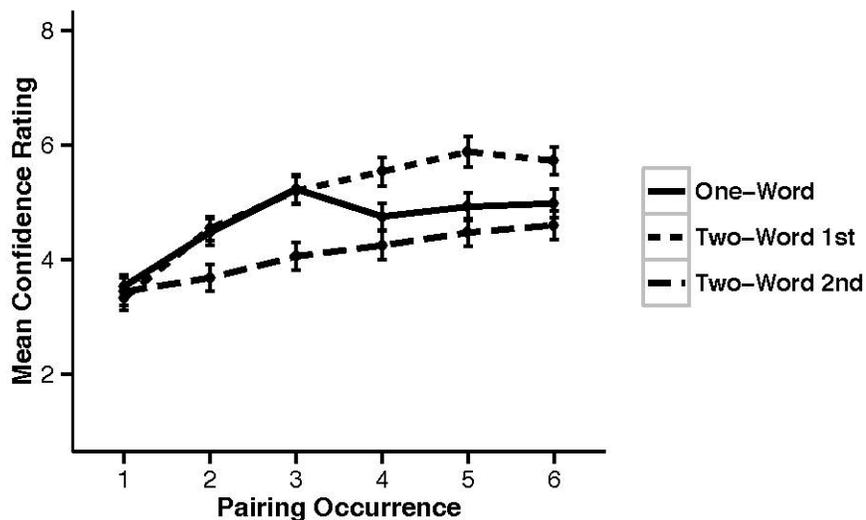
Participants successfully learned both One- and Two-Word pairings [One:  $t(94) = 11.39$ ,  $p < 0.001$ , 95% CI [0.47, 0.57],  $d = 1.17$ ; Both:  $t(94) = 5.29$ ,  $p < 0.001$ , 95% CI [0.12, 0.19],  $d = 0.54$ ]. As in Experiment 1, participants were more likely to learn one label for an object than two [ $t(94) = 11.16$ ,  $p < 0.001$ , 95% CI [0.3, 0.43],  $d = 1.83$ ]. However, there was no relation in the learning of corresponding Two-Word 1<sup>st</sup> and 2<sup>nd</sup> pairings. A mixed effects logistic regression analysis (with random effect terms of subject and item) analyzing the individual test trials revealed that Two-Word 1<sup>st</sup> pairings had no

relation to Two-Word 2<sup>nd</sup> pairings (Beta = -0.28, odds ratio = 0.75, STE = 0.21,  $p = 0.19$ , 95% CI [-0.70, 0.14]), contrary to what we found in Experiment 1. Blocked training of the two labels for an object eliminated item-level competition. Thus, the evidence to this point favors the independent learning of competing pairings when separated in time.

Are participants aware of the cross-block introduction of competing items?

Analyses of the confidence ratings suggests that they might be. Figure 6 displays the ratings results across each word-object co-occurrence. A repeated measures ANOVA (by-subject) revealed a significant effect of Occurrence [ $F(5, 470) = 67.28$ ,  $p < 0.001$ , 95% CI [0.35, 0.47],  $\eta^2_G = 0.06$ ], Pairing Type [ $F(2, 188) = 22.75$ ,  $p < 0.001$ , 95% CI [0.1, 0.3],  $\eta^2_G = 0.03$ ], and an interaction between the two [ $F(10, 940) = 11.66$ ,  $p < 0.001$ , 95% CI [0.07, 0.14],  $\eta^2_G = 0.01$ ; a by-item ANOVA yielded the same pattern of results]. Independent samples t-tests (with Bonferonni corrections) showed that Two-Word 2<sup>nd</sup> pairings were rated lower than One-Word pairings [ $t(94) = 4.29$ ,  $p < 0.001$ , 98.3% CI [0.25, 0.89],  $d = 0.27$ ] and Two-Word 1<sup>st</sup> pairings [ $t(94) = 5.24$ ,  $p < 0.001$ , 98.3% CI [0.51, 1.4],  $d = 0.46$ ], demonstrating that overlapping statistics decreased adults' confidence in their knowledge of those pairings. The presence of a second label for Two-Word objects in Block 2 also lowered participants' confidence that they knew the One-Word pairings. One-Word pairings were rated lower than Two-Word 1<sup>st</sup> pairings [ $t(94) = -3.83$ ,  $p < 0.001$ , 98.3% CI [-0.63, -0.14],  $d = 0.19$ ]. The low confidence ratings for One-word pairings was largely driven by the second block of training, as suggested by the Occurrence by Pairing Type interaction. One-Word pairings were presented in Block 1 and also in Block 2, at which point new, overlapping statistics were introduced. This change in the presence of overlapping statistics across blocks is reflected in the

confidence ratings. Although at occurrence 3 (Block 1), One- and Two-Word 1<sup>st</sup> pairings were no different, the fourth time a One-Word pairing occurred (Block 2), confidence decreased. Two-Word 1<sup>st</sup> pairings, on the other hand, were rated highest in confidence because they were only presented in Block 1, when no overlapping statistics had been encountered. Thus, when encountered, overlapping word-object pairings reduced confidence for all pairings. Although this pattern is consistent with the possibility that participants were aware of the overlapping statistics, this does not have to be the case. Participants could just have known that they did not know the Two-Word 2<sup>nd</sup> items well. Although participants' confidence ratings showed One-Word versus Two-Word effects in Experiment 2, where they did not in Experiment 1, one aspect of the confidence ratings is the same: The confidence ratings did not predict performance (see Table 3). If participants do indeed have some awareness of the presence of overlapping statistics, this awareness may not be directly linked to the presence of competition.



**Figure 6.** Mean ratings (and standard errors of the mean) for One- and Two-Word objects for each occurrence in training in Experiment 2, collapsed across Language group. Participants were asked to rate, from 1-10, their knowledge of the word for each object (1 meaning I do not know the name for this object, and 10 meaning I do know the name for this object).

In conclusion, Experiment 2 showed that inter-item competition is reduced (perhaps even eliminated) when the overlapping pairings are blocked in time. Note that although we found no evidence of item-level competition, our results include patterns that have in the past been taken as indicative of competition effects (see Kachergis et al., 2012; Ichinco et al., 2009; Yurovsky et al., 2013): One-Word pairings were learned significantly better than Two-Word pairings, and adults were more likely to learn one label than both labels for an object. Altogether, these results point to the role of time as a mediator of competition: the separation in time of overlapping statistics reduces item-level competition. Also, as in Experiment 1, we found no relation between subjects' confidence ratings nor their language history and performance in learning the overlapping pairings. These additional findings could mean that competition effects may emerge from implicit learning machinery.

### **Experiment 3**

Is the finding in Experiment 2 really about time? Or is it about the role of a contextual cue that segregates competing co-occurrence data? When one attempts to learn a second language, for example, the words from the first and second languages are segregated by time, context, and by their phonological properties which differ between the two languages. If one were to attempt to learn two languages *at the same time*, the phonological properties of the two languages may be sufficient in and of themselves to reduce competition. Accordingly, to better understand the constraints on inter-item competition, Experiment 3 used the same inter-mixed presentation of Two-Word items that was used in Experiment 1, but distinguished the two words for the same object, not

by time, but by a distinct property of the word. Specifically, the labels were structured so that all labels for One-Word pairings, and *one* label for each Two-Word object were disyllabic and ended in a vowel. The second label for each Two-Word object, however, was monosyllabic and ended with the stop consonant /k/. We chose this contextual cue for two reasons. First, there is ample evidence that the phonological structure of words (including syllable length and phonotactics) aids in resolving ambiguities in word and sentence meanings (Au & Glusman, 1990; Durieux & Gillis, 2001; Kelly, 1996). Additionally, learners are successful at using phonological cues and word structure in statistical learning (Escudero et al., 2015; Lew-Williams & Saffran, 2012; Sahni, Seidenberg, & Saffran, 2010).

## Method

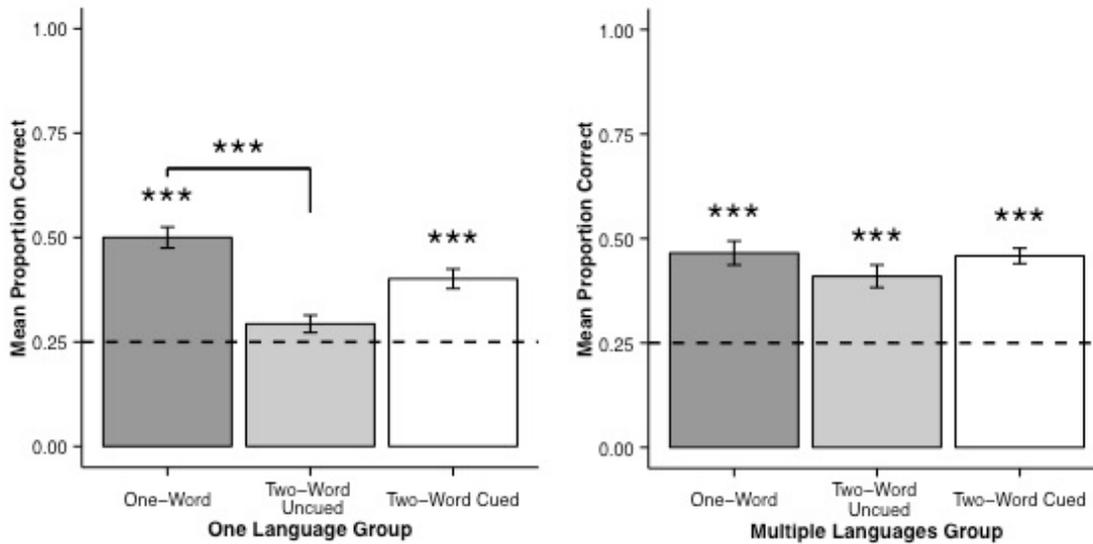
**Participants.** Participants were 102 adults that did not participate in Experiment 1 or 2 and who were recruited through the same procedures as in the earlier experiments: 54 (14 males) participants were grouped into the One Language group and 48 (21 males) into the Multiple Languages group. There were no significant differences in the demographic factors between the two groups other than country of birth (see Table 2).

**Stimuli and Design.** All aspects of the stimuli and design of Experiment 3 were the same as Experiment 1 except for the words presented to participants. Twelve unique, novel, English phonotactic pseudowords were generated, organized by their syllable length and endings. All six words for One-Word pairings were two syllables in length and ended in a vowel. One of the words for each Two-Word object, a total of six, also followed this structure (Two-Word Uncued). The other word for each Two-Word object, a total of six, was a unique one-syllable word that ended in the stop consonant /k/ (Two-

Word Cued; see Figure 2). Note that, unlike Experiment 1, all words in Experiment 3 followed some kind of structure (either mono-syllabic with a vowel ending, or di-syllabic with a /k/ ending). The key manipulation here, however, is that the second labels of the Two-Word objects followed a different unique structure than all other words, with the goal that this unique structure would highlight these pairings. As in Experiment 1, for the Two-Word pairings, there were on average 1.77 repetitions of an individual word before the other word for that same object was presented.

### **Results and discussion**

Figure 7 shows the pattern of overall performance separately for the One Language and Multiple Languages groups, because in this experiment, self-reported language history mattered. Participants' proportion correct at test for the One-Word, Two-Word Uncued, and Two-Word Cued pairings were submitted to an ANOVA for a 3 (Pairing Type) by 2 (Language Group) mixed design. The by-subject analysis revealed a main effect of Pairing Type [ $F(2, 200) = 10.02, p < 0.001, 95\% \text{ CI } [0.03, 0.17], \eta^2_G = 0.05$ ; the by-item analysis showed the same result]. Bonferonni corrected tests showed that performance on Two-Word Uncued pairings ( $M = 0.35, SD = 0.25$ ) was significantly lower than performance on One-Word pairings [ $M = 0.48, SD = 0.27$ ],  $t(101) = 4.27, p < 0.001, 98.3\% \text{ CI } [0.06-0.2], d = 0.52$ ] and Two-Word Cued pairings [ $M = 0.43, SD = 0.22$ ],  $t(101) = 2.69, p = 0.03, 98.3\% \text{ CI } [0.008-0.15], d = 0.35$ ]. There was no difference between One-Word and Two-Word Cued pairings [ $t(101) = 1.81, p = 0.22$ ]. Cueing one of the associated words for the Two-Word pairings thus facilitated the learning of those pairings.

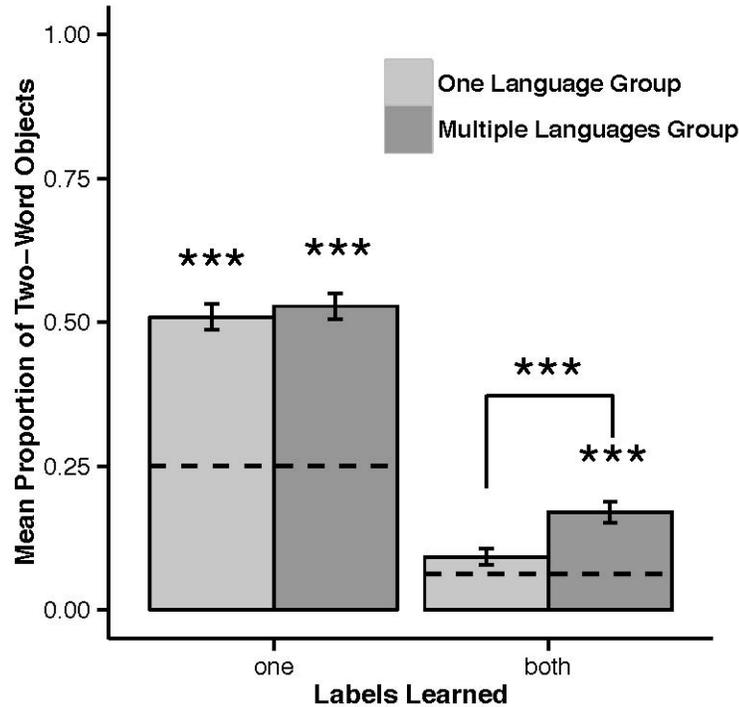


**Figure 7.** Means and standard errors of the mean for accuracy for One- and Two-Word pairings for a) One Language Group, and b) Multiple Languages Group for Experiment 3, when second labels (Two-word Cued) were highlighted using a linguistic cue. Dashed lines denote chance performance and asterisks denote significant differences at the  $p < 0.001$  level.

In contrast, to Experiments 1 and 2, Language Group had an effect on learning. The by-subject analysis showed a reliable Language Group by Pairing Type interaction [ $F(2, 200) = 3.12$ ,  $p = 0.046$ , 95% CI [0.0003, 0.08],  $\eta^2_G = 0.02$ ; the by-item analysis showed a main effect of Language Group:  $F(1,22) = 4.87$ ,  $p = 0.04$ , 95% CI [0.006, 0.43],  $\eta^2_G = 0.05$ ]. Pairwise comparisons (with Bonferonni corrections) demonstrated that the One Language group's performance on both One-Word and Two-Word Cued pairings were better than Two-Word Uncued pairings [One-Word vs. Two-Word Uncued:  $t(53) = 5.57$ ,  $p < 0.001$ , 99.17% CI [0.12, 0.31],  $d = 0.89$ ; Two-Word Cued vs. Two-Word Uncued:  $t(53) = 2.87$ ,  $p = 0.036$ , 99.17% CI [0.005, 0.21],  $d = 0.48$ ], with no difference between One-Word and Two-Word Cued pairings [ $t(53) = 2.34$ ,  $p = 0.14$ ]. However, for participants in the Multiple Languages group, there were no significant differences

between any of the pairing types [One-Word vs. Two-Word Uncued,  $t(47) = 1.10$ ,  $p = 1.0$ ; One-Word vs. Two-Word Cued,  $t(47) = 0.16$ ,  $p = 1.0$ ; and Two-Word Uncued vs. Two-Word Cued,  $t(47) = -1.03$ ,  $p = 1.0$ ]. In brief, the distinct phonotactic structure for a subset of words had different effects depending on the language history of the participants. For participants who self-reported experiences with learning multiple languages, both sets of overlapping word-object pairings were as learnable as the One-Word pairings. This was not the case for participants who identified as only speaking English. These language differences are emphasized by considering participants' performance relative to chance. Participants in the Multiple Languages group performed above chance on all pairing types (One-Word,  $t(47) = 5.15$ ,  $p < 0.001$ , 95% CI [0.38, 0.55],  $d = 0.74$ ; Two-Word Uncued,  $t(47) = 4.06$ ,  $p < 0.001$ , 95% CI [0.33, 0.49],  $d = 0.59$ ; Two-Word Cued,  $t(47) = 7.73$ ,  $p < 0.001$ , 95% CI [0.4, 0.51],  $d = 0.64$ ). Participants in the One Language group performed above chance only on One-Word pairings [ $t(53) = 7.27$ ,  $p < 0.001$ , 95% CI [0.43, 0.57],  $d = 0.99$ ] and Two-Word Cued pairings [ $t(53) = 4.70$ ,  $p < 0.001$ , 95% CI [0.34, 0.47],  $d = 0.64$ ]. The One Language group did not learn Two-Word Uncued pairings [ $t(53) = 1.51$ ,  $p = 0.14$ ]. These group differences do not appear to be due to a subset of the Multiple Languages participants. In separate ANOVAS for subgroups of the Multiple Languages group (defined as in Experiment 1: Native ( $N = 21$ ) versus Non-native proficiency ( $N = 27$ ), Early ( $N = 22$ ) versus Late ( $N = 25$ ) age of exposure, and Native English speakers ( $N = 27$ ) versus Non-native English speakers ( $N = 21$ )), there were no reliable differences. In brief, the task of having tried to learn with at least some moderate success a second language appears to be the principle discriminating factor.

Figure 8 shows the breakdown of performance for Two-Word pairings for the two language groups. For both groups of participants, the proportion of Two-Word objects for which one label was learned was above chance [One Language group:  $t(53) = 8.46$ ,  $p < 0.001$ , 95% CI [0.45, 0.57],  $d = 1.15$ ; Multiple Languages group:  $t(47) = 8.50$ ,  $p < 0.001$ , 95% CI [0.46, 0.59],  $d = 1.23$ ) and did not differ between the groups [ $t(100) = 0.41$ ,  $p = 0.68$ ]. However, the learning of both labels for Two-Word objects was above chance in the Multiple Languages group [ $t(47) = 4.13$ ,  $p < 0.001$ , 95% CI [0.12, 0.22],  $d = 0.60$ ] and was higher than participants in the One-Language group [ $t(100) = 2.44$ ,  $p = 0.02$ , 95% CI [0.015, 0.14],  $d = 0.48$ ), who did not reach above chance levels of performance [ $t(53) = 1.58$ ,  $p = 0.12$ ]. Similar to both Experiment 1 and 2, however, for both groups, the proportion of objects for which only one label was learned was much higher than for objects for which both labels were learned which *could* be due to inter-item competition in both groups. Alternatively, the overall better performance of the Multiple Languages group relative to the One Language group on Two Word pairings could mean reduced inter-item competition in the Multiple Languages group.

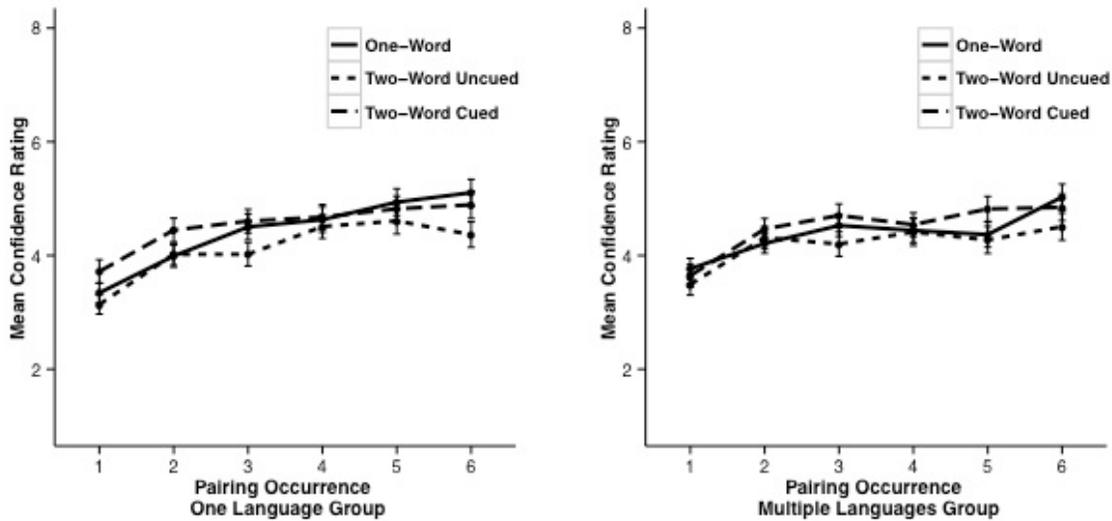


**Figure 8.** Means and standard errors of the mean for the proportion of Two-Word objects for which one label or both labels were learned for the One Language and Multiple Languages groups in Experiment 3, when second labels were highlighted with a linguistic cue. Dashed lines denote chance performance and asterisks denote significant differences at the  $p < 0.001$  level.

A logistic regression (with random effect terms of subject and item) examining how accuracy for Two-Word Cued pairings predicted accuracy for the corresponding Two-Word Uncued pairings at the individual test trial level showed a negative and significant effect in the One Language group (Beta = -0.86, odds ratio = 0.42, STE = 0.39,  $p = 0.03$ , 95% CI [-1.62, -0.10]), a pattern clearly implicating inter-item competition. The participants in the Multiple Languages group, however, did not show reliable dependencies between learning the overlapping items (Beta = -0.29, STE = 0.28,  $p = 0.30$ ). For the One Language group, accuracy on Two-Word Cued tests decreased the

likelihood of accuracy on Two-Word Uncued tests, therefore also reducing the number of pairings for which they learned both words. For participants in the Multiple Languages group, there was no item-by-item relation between the two types of pairings.

The analyses of the trial-by-trial confidence ratings, shown in Figure 9, indicate a complicated set of relations between these ratings and participants' learning performance. A mixed measures ANOVA (by-subjects) revealed a main effect of Occurrence ( $F(5, 500) = 24.51, p < 0.001, 95\% \text{ CI } [0.13, 0.25], \eta^2_G = 0.04$ ), Pairing Type ( $F(2, 200) = 5.64, p = 0.004, 95\% \text{ CI } [0.006, 0.12], \eta^2_G = 0.005$ ), and an Occurrence by Pairing Type interaction ( $F(10, 1000) = 2.35, p = 0.01, 95\% \text{ CI } [0.001, 0.03], \eta^2_G = 0.002$ ). The main effect of Language Group, all interactions with Language Group, and the three-way interaction, were not significant (the by-subject analysis overall patterned the same except for a significant Language Group by Occurrence interaction:  $F(10, 220) = 3.39, p = 0.007, 95\% \text{ CI } [0.01, 0.22], \eta^2_G = 0.003$ ). Bonferroni corrected tests showed that for both groups, ratings for Two-Word Cued pairings ( $M = 4.52, SD = 1.83$ ) were higher than Two-Word Uncued pairings [ $M = 4.15, SD = 1.87, t(101) = 3.24, p = 0.005, 98.3\% \text{ CI } [0.09, 0.64], d = 0.2$ ], and there was no difference between Two-Word Cued pairings and One-Word pairings [ $M = 4.41, SD = 1.83, t(101) = 1.09, p = 0.84$ ]. One-Word pairings were not different than Two-Word Uncued pairings ( $t(101) = 2.16, p = 0.1$ ). This pattern suggests that both groups of learners noticed the Two-Word Cued words with their distinctive phonotactic properties relative to the other words in the training set and thus thought they knew their referents better than the referents of the other words. By the last occurrence of a pairing, only confidence ratings for the Multiple Languages group for some of the pairings were significantly predictive of performance at test (see Table 3).



**Figure 9.** Mean ratings (and standard errors of the mean) for One- and Two-Word objects for each occurrence in training for a) One Language group and b) Multiple Languages Group in Experiment 3. Participants were asked to rate, from 1-10, their knowledge of the word for each object (1 meaning I do not know the name for this object, and 10 meaning I do know the name for this object).

The findings from this experiment make a new contribution to understanding inter-item competition in statistical word-referent learning. They show that there exists combinations of cues and learners such that overlapping pairings –even when interleaved in time – may be learned independently. Therefore, time may not be the only factor that limits competition. In addition, the findings tell us that there are individual differences, potentially linked to language learning history, that matter for how inter-item competition plays out. There are several unresolved hypotheses about the specific nature of the effect: these differences could be due to experience with multiple languages, experience with languages other than English, or individual properties that lead people to attempt to learn and to self-report more success at learning languages.

### General Discussion

The experiments were designed to address two empirical questions about cross-situational word-referent learning: First, is there competition at the item level? This is a core idea in many models of cross-situational learning but one that has not been directly demonstrated and about which there is some debate (Kachergis et al., 2012; K. Smith et al., 2011; Yu & Smith, 2012). Second, what are the temporal and contextual limits on item-level competition? Answers to this second question are relevant to determining the mechanism(s) through which competition occurs. The findings from Experiment 1 and from the One Language group of Experiment 3 provide clear evidence of competition at the item level in that the likelihood of a participants' knowledge of two competing associations with the same input statistics were negatively related to each other. The results from Experiment 2 and from the Multiple Languages group of Experiment 3 show that item-level competition does not always occur. Further, the results from Experiments 2 and 3 considered jointly indicate that the temporal separation of competing items may be sufficient to eliminate competition (Poepsel & Weiss, 2014; Yurovsky et al., 2013) but it is not necessary. Some participants, those who reported experiences learning multiple languages, showed independent learning of competing items intermixed in time when the two competing labels were phonotactically distinct. In sum, the role of competition between overlapping associations in statistical word-referent learning depends on the temporal window in which competing associations are encountered, the distinctiveness of the competing associations, and the properties of the individual learner.

Is there a unified set of principles under which the circumstances of inter-item competition can be explained? It is certainly possible that the answer to this question is

“no.” There are many components to cross-situational word-referent learning including the auditory processing and representation of the labels, the visual perception of the objects, as well as associated attention and memory processes (Smith, Suanda, & Yu, 2014). What we already know is that each of the processes is sensitive to and adapts to the statistical regularities in the learning environment (e.g., Saffran & Thiessen, 2007; Smith et al., 2014). Those different statistical learning systems – speech processing, visual attention, memory -- could have different operating characteristics that determine how and when competition occurs (Conway & Christiansen, 2006; Frost, Armstrong, Siegelman, & Christiansen, 2015). Thus, the temporal separation of competing associations could limit competition through different mechanisms and for a different reason than phonologically distinct labels do.

However, a general set of principles – and ones that may clarify the contexts in which competition is likely to occur – may also apply across these different statistical learning components. In order for an association from prior learning to compete –and potentially inhibit-- the learning of a new co-occurrence, that prior association must be activated. Temporal proximity and similarity are key factors that determine whether current input activates memories for some prior input (e.g. Estes, 1986). Under these principles, inter-item competition was high in Experiment 1 and low in Experiment 2 because the temporal proximity of highly similar overlapping associations in Experiment 1 led to competitive activation, but their temporal separation in Experiment 2 limited the ability of currently present co-occurrences to co-activate the overlapping item in the competitive process. Under these principles, the finding of reduced competition given temporally intermixed overlapping associations for the Multiple Languages participants

in Experiment 3 is also straightforward: the distinctiveness of the two words associated with the same object limited co-activation and inhibition.

Can these general principles also explain why distinctiveness did not mitigate competition for the One Language participants in Experiment 3 but instead strengthened that competition in one direction, with the more unique labels inhibiting the learning of competing labels for the same object? One possibility is that the uniqueness of the monosyllabic /k/-ending labels relative to the other to-be-learned labels attracted the attention of the One Language participants early in learning (Hunt, 1995; Schmidt, 1991; Wallace, 1965) so that these associations increased in strength more rapidly than their competitors and therefore effectively inhibited those competitors. To explain the whole test performance pattern, we have to assume that the less salient competitor, when encountered, activated the more salient one for the One Language participants, but not for the Multiple Languages participants. The Multiple Languages participants, by hypothesis, may have consistently tracked both the words in the Two-Word pairings but with competition reduced because of their increased discriminability. Clearly, these are conjectures that need further empirical tests. However, determining the role of timing and item similarity on how competition plays out – and the role of the history of the learner – are clearly critical next steps to understanding cross-situational word-referent learning (Smith, et al., 2014; Yu & Smith, 2012).

The present experiments examined competition in the specific context of aggregating word-referent associations across individually ambiguous trials in order to discover the underlying set of word-object correspondences. Competition between overlapping associations has been viewed as especially important to this form of learning

because such competition could hasten learning by removing weaker spurious associations (e.g., Yu & Smith, 2012). The results from Experiments 1 and 2 suggest that in this learning task, competition is a temporally local phenomenon *between relatively weak associations*, conclusions that may not apply to all forms of competition (see Dumay & Gaskell, 2007). In Experiment 1, overlapping pairings were separated on average by 1.77 repetitions of each word, a moderately short period for word-referent associations that were being slowly formed. In Experiment 2, presenting all 6 occurrences of the first word before the second word eliminated this competition. How much prior learning is needed to eliminate competition between overlapping items? Several studies have manipulated the spacing between repetitions of individual word-object pairings in the cross-situational word learning paradigm and shown effects on learning (Romberg & Yu, 2015; Vlach & Johnson, 2013). The same manipulations over different time periods for measures of inter-item competition would help to determine the time constraints and time course of inter-item competition.

A growing set of results suggests that cross-situational word-referent learning may be a form of implicit statistical learning, including results suggesting that explicit trial-by-trial tests of learning disrupt the discovery of the underlying word-referent correspondences. The lack of systematic relations between participants' trial-by-trial confidence judgments and actual learning performance in our experiments is consistent with this conclusion. Critically, lower confidence –and apparent awareness of competitors – also did not predict item-level competition. If statistical word-referent learning is like the learning of patterns of co-occurrences and predictive relations in other domains (Conway & Christiansen, 2006; Kim, Seitz, Feenstra, & Shams, 2009; Reber,

1967; Turk-Browne, Jungé, & Scholl, 2005; Turk-Browne, Scholl, Chun, & Johnson, 2009), then learners' explicit strategies and hypotheses may not be straightforwardly linked to, nor predictive of, their learning progress (see Perruchet & Pacton, 2006). However, strong conclusions are not yet warranted as explicit strategies may also sometimes benefit learning. Poepsel & Weiss (2014) showed that an explicit contextual cue pointing to the presence of multiple words for a referent (speaker voice or explicit instructions) boosted confidence ratings in a cross-situational word-learning task.

The present results also do not provide a clear indication of the relevant participant differences that led to the different performance patterns of the One Language and Multiple Languages groups in Experiment 3. All participants were recruited at a U.S. university. All participants spoke English. They differed as to whether they also self-indicated as speaking an additional language. Some of the participants in the Multiple Languages group were native speakers of English, some were not, and they varied markedly in their experience with a second language from using two languages from very early in childhood to learning a second language in college. When we examined if the extent of second language experience affected performance, partitioning participants in the Multiple Languages group by proficiency in the second language, age of exposure to the second language, or whether English was the native language, yielded no reliable differences in performance in Experiment 3 (nor Experiments 1 and 2).

Overall, our results suggest that the observed differences with respect to language history do not depend on being bilingual in the strict sense, but may reflect effects of moderate experiences with other languages or language processing abilities that support attempts to learn other languages. Although the word-structure differences and

competition in Experiment 3 may be conceptualized as being similar to the context of learning a second language or to a bilingual environment (multiple phonologically distinct labels for the same object; Hernandez, Li, & MacWhinney, 2005), and although bilingual children have been reported to show less competition among overlapping words in some word learning tasks (Bialystok, Barac, Blaye, & Poulin-Dubois, 2010; Byers-Heinlein & Werker, 2009, 2013; Davidson & Tell, 2005; Houston-Price, Caloghiris, & Raviglione, 2010; Yoshida et al., 2011), the unexpected findings in Experiment 3 are probably best not interpreted as a “bilingual effect” on statistical word-referent learning. They could instead be due to different aptitudes in language processing in the two groups, to modest experiences with different languages, and/or to more English-centric versus less-English centric biases in mapping words to referents. Thus, the source of the group effects in Experiment 3 should be taken as preliminary. However, the findings clearly indicate that there are relevant individual differences with respect to competition in word-referent learning that may be related to language learning history in some way. Specifying the nature of these individual differences will be critical to understanding competitive processes and how –and when— they support statistical learning.

Inter-item competition plays a central role in current theories and debates about lexical learning by adults and by infants (MacWhinney, 1989; McClelland & Elman, 1986; McMurray et al., 2012) because in a noisy learning environment with many spurious co-occurrences, competition provides a way of cleaning-up the co-occurrence data. However, a learning system that learns too rapidly, settling on initially strong co-occurrences, runs the risk of learning the wrong regularities, including not learning that some objects have multiple labels. A learning system that treats all co-occurrences as

part of the same big data set, rather than partitioning them into distinct and non-interacting sets, runs the risk of not learning anything or not being able to find the different latent structure in multiple domains (such as two different languages; e.g., Quian, Jaeger, & Aslin, 2012). Understanding how timing, context, and item distinctiveness help learners solve these fundamental problems is essential to understanding how we learn from ambiguous data.

These issues and their solutions are also intertwined in current debates as to whether lexical learning is best understood as a form of hypothesis testing or associative learning (Kachergis et al., 2012; Medina, Snedeker, Trueswell, & Gleitman, 2010; Romberg & Yu, 2014; Trueswell, Medina, Hafri, & Gleitman, 2013; Yu & Smith, 2012; Yurovsky, Smith, & Yu, 2013; see Yurovsky & Frank, 2015 for an integrative account). A central idea within many variants of hypothesis-testing accounts is that learners represent specific hypotheses about word–referent pairings and then, in the face of experienced evidence, select among those hypotheses based on some principled inference procedure (Frank et al., 2009; Halberda, 2006; Siskind, 1996; Trueswell et al., 2013). One common inferential principle is mutual-exclusivity (Markman, 1990; Merriman et al., 1989), the constraint of forming and confirming hypotheses in which referents have a single label. Within associative theories, statistical learning emerges from the strengthening and weakening of associations between words and referents as a function of co-occurrence strength and competition among associations (Smith et al., 2014; Suanda & Namy, 2012; Yu & Smith, 2007, 2012; Yurovsky et al., 2013). Yu and Smith (2012) have argued that both classes of theories might be understood as emerging from the same attentional and memorial processes and differ primarily in the parameters

concerning how many co-occurrences are registered within a single learning event and the degree of competition among overlapping associations, with hypothesis testing characterizing the extremes of selectivity and winner-take-all competitive processes. These issues have led to a now large literature of mixed evidence showing that participants sometimes do and sometimes do not learn overlapping associations, with the former taken as supportive of hypothesis testing (Frank et al., 2009; Medina, et al., 2010; Siskind, 1996; Trueswell et al., 2013) and the latter as supportive of associative learning (Smith et al., 2014; Suanda & Namy, 2012; Yu & Smith, 2007, 2012; Yurovsky et al., 2013).

The present results contribute to these debates by providing a direct test of inter-item competition. The contributions of the results are these: Inter-item competition is a central process in cross-situational word learning. The degree to which one word for a referent is learned is negatively related to the learning of another word for that referent. This inter-item competition is constrained by timing, association distinctiveness, and language learning history. The finding that learning sets that are separated in time do not compete implicates competitive processes with limited dynamic windows. However, the findings that association distinctiveness and language learning history also modulate competition implicate effects of past learning on these transient competitive processes. Competition has been proposed as a critical process in “cleaning” the data and speeding statistical word-referent learning. The present experiments provide the first direct experimental evidence for inter-item competition.

### **Acknowledgements**

This research was supported by NSF Graduate Research Fellowship DGE-1004163 awarded to VLB, NIH NRSA F32HD075577 awarded to DY, and in part by NICHD grant RO1 HD056029 to LBS (PI Chen Yu). We would like to thank Mary Clare Charles, Eden Faye, and Rachel Rapsinski for help with recruiting and running participants.

References

- Adesope, O.O., Lavin, T., Thompson, T., & Ungerleider, C. (2010). A systematic review and meta-analysis of the cognitive correlates of bilingualism. *Review of Educational Research, 80*(2), 207-245.
- Alloppenna, P.D., Magnuson, J.S., & Tanenhouse, M.K. (1998). Tracking the time course of spoken word recognition using movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419-439.
- Au, T. K., & Glusman, M. (1990). The principle of mutual exclusivity in word learning: To honor or not to honor? *Child Development, 61*, 1474-1490.
- Bartolotti, J., Marian, V., Schroeder, S.R., & Shook, A. (2011). Bilingualism and inhibitory control influence statistical learning of novel word forms. *Frontiers in Psychology, 2*(324), doi: 103389/fpsyg.2011.00324.
- Bates, E., & MacWhinney, B. (1989). Competition, variation, and language learning. In B. MacWhinney (Ed.) *Mechanisms of Language Learning* (pp. 157-193). Hillsdale, NJ: Erlbaum.
- Bialystok, E., Barac, Raluca, Blaye, A., & Poulin-Dubois, D. (2010). Word mapping and executive functioning in young monolingual and bilingual children. *Journal of Cognitive Development, 11*(4), 485-508.
- Blythe, R.A., Smith, K., & Smith, A.D.M. (2010). Learning times for large lexicons through cross-situational learning. *Cognitive Science, 34*, 620-642.
- Brojde, C.L., Ahmed, S., & Colunga, E. (2012). Bilingual and monolingual children attend to different cues when learning new words. *Frontiers in Psychology, 3*(155), doi: 10.3389/fpsyg.2012.00155

- Byers-Heinlein, K., & Werker, J.F. (2009). Monolingual, bilingual, trilingual: infants' language experience influences the development of a word-learning heuristic. *Developmental Science, 12*(5), 815-823.
- Byers-Heinlein, K., & Werker, J.F. (2013). Lexicon structure and the disambiguation of novel words: Evidence from bilingual infants. *Cognition, 128*, 407-416.
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*, 354–380.
- Conway, C.M., & Christiansen, M.H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological Science, 17*(10), 905-912.
- Cutler, A. (1995). Spoken word recognition and production. In J. L. Miller & P. D. Eimas (Eds.), *Handbook of perception and cognition: Vol. 11. Speech, language, and communication* (pp. 97–136). San Diego: Academic Press.
- Davidson, D., & Tell, D. (2005). Monolingual and bilingual children's use of mutual exclusivity in the naming of whole objects. *Journal of Experimental Child Psychology, 92*. 25-45.
- Desimone, R. & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193-222.
- Dumay, N., & Gaskell, M.G. (2007). Sleep-associated changes in the mental representation of spoken words. *Psychological Science, 18*(1), 35-39.
- Durieux, G., & Gillis, S. (2001). Predicting grammatical classes from phonological cues. In Jürgen Weissenborn, and Barbara Höhle (Eds.), *Approaches to bootstrapping:*

*Vol. 1. Phonological, lexical, syntactic and neurophysiological aspects of early language acquisition* (pp. 189-229). Amsterdam: John Benjamins.

Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology*.

New York, NY: Teachers College, Columbia University.

Elman, J., Hare, M., & McRae, K. (2005). Cues, constraints, and competition in sentence processing. In M. Tomasello and D. Slobin (Eds.) *Beyond Nature-Nurture: Essays in Honor of Elizabeth Bates* (pp. 111-138). Mahwah, NJ: Lawrence Erlbaum Associates.

Escudero, P., Mulak, K.E., & Vlach, H.A. (2015). Cross-situational learning of minimal word pairs. *Cognitive Science*. Advance online publication. doi: 10.1111/cogs.12243

Estes, W. K. (1986). Memory storage and retrieval processes in category learning. *Journal of Experimental Psychology: General*, 115(2), 155-174.

Fitneva, S. A., & Christiansen, M. H. (in press). Developmental changes in cross-situational word learning: The inverse effect of initial accuracy. *Cognitive Science*.

Fitneva, S. A., & Christiansen, M. H. (2011). Looking in the wrong direction correlates with more accurate word learning. *Cognitive Science*, 35(2), 367-380

Frank, M.C., Goodman, N.D., & Tenenbaum, J.B. (2009). Using speakers' referential intentions to model early cross-situational word learning. *Psychological Science*, 20(5), 578-585.

Frost, R., Armstrong, B.C., Siegelman, N., & Christiansen, M.H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends*

- in Cognitive Sciences*, 19(3), 117-125.
- Halberda, J. (2006). Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology*, 53(4), 310-344.
- Hernandez, A., Li, P., & MacWhinney, B. (2005). The emergence of competing modules in bilingualism. *Trends in Cognitive Sciences*, 9(5), 220-225.
- Hilchey, M.D., Klein, R.M. (2011). Are there bilingual advantages on nonlinguistic interference tasks? Implications for the plasticity of executive control processes. *Psychonomic Bulletin & Review*, 18, 625-658.
- Horst, J.S., Scott, E.J., & Pollard, J.A. (2010). The role of competition in word learning via referent selection. *Developmental Science*, 13(5), 706-713.
- Houston-Price, C., Caloghiris, Z., & Raviglione, E. (2010). Language experience shapes the development of the mutual exclusivity bias. *Infancy*, 15(2), 125-150.
- Howard, D., Nickels, L., Coltheart, M., & Cole-Virtue, J. (2006). Cumulative semantic inhibition in picture naming: Experimental and computational studies. *Cognition*, 100, 464-482.
- Hunt, R.R. (1995). The subtlety of distinctiveness: What von Restorff really did. *Psychonomic Bulletin & Review*, 2(1), 105-112.
- Ichinco, D., Frank, M.C., & Saxe, R. (2009). Cross-situational word learning respects mutual exclusivity. Proceedings of the 31<sup>st</sup> Cognitive Science Conference, 31. Austin, TX: Cognitive Science Society.

- Kachergis, G., Yu, C., & Shiffrin, R.M. (2012). An associative model of adaptive inference for learning word-referent mappings. *Psychonomic Bulletin and Review*, *19*, 317-324.
- Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, *16*(4), 705-710.
- Kelly, M.H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignment. *Psychological Review*, *99*(2), 349-364.
- Kim, R., Seitz, A., Feenstra, H., & Shams, L. (2009). Testing assumptions of statistical learning: is it long-term and implicit? *Neuroscience letters*, *461*(2), 145-149.
- Kroll, J.F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*(5), 497-514.
- Levelt, W.J.M., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, *22*, 1-75.
- Lew-Williams, C., & Saffran, J.R. (2012). All words are not created equal: Expectations about word length guide infant statistical learning. *Cognition*, *2012*, 241-246.
- Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior Research Methods*, *28*(2), 202-210.
- Markman, E.M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14* (1), 57-77.
- Marslen-Wilson, W. D. (1990). Activation, competition and frequency in lexical access. In G. T. M. Altmann (Ed.), *Cognitive models of speech processing*:

- Psycholinguistic and computational perspectives* (pp. 148-172). Cambridge, MA: MIT Press.
- MacWhinney, B. (1989). Competition and lexical categorization. In R. Corrigan, F. Eckman, & M. Noonan (Eds.), *Linguistic categorization* (pp. 195-242). New York: John Benjamins.
- McClelland, J.L. & Elman, J.F. (1986). The TRACE model of speech perception. *Cognitive Psychology*, 18, 1-86.
- McMurray, B., Horst, J. S., & Samuelson, L.K. (2012). Word learning emerges from the interaction of online referent selection and slow associative learning. *Psychological Review*, 119(4), 831-877.
- McRae, K., Spivey-Knowlton, M.J., & Tanenhaus, M.K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, 38, 283-312. *Proceedings of the National Academy of Sciences*, 108, 9014-9019.
- Medina, T.N., Snedeker, J., Trueswell, J.C., & Gleitman, L.R. (2010). How words can and cannot be learned by observation.
- Mensink, G.J. M. & Raaijmakers, J.G.W. (1988). A model for interference and forgetting. *Psychological Review*, 95(4), 434-455.
- Merriman, W.E. (1999). Competition, attention, and young children's lexical processing. In B. MacWhinney (Ed.), *The emergence of language* (pp. 331-358). Mahwah, NJ: Erlbaum.

- Merriman, W.E., Bowman, L.L., & MacWhinney, B. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development, 54*(3/4), 1-129.
- Norman, K.A., Newman, E.L. & Detre, G.J. (2007). A neural network model of retrieval-induced forgetting. *Psychological Review, 114*(4), 887-953.
- Oppenheim, G.M., Dell, G.S., & Schwartz, M.F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition, 114*(2), 227-252.
- Poepsel, T.J. & Weiss, D.J. (2014). Context influences conscious appraisal of cross-situational statistical learning. *Frontiers in Psychology, 5*(691), doi: 10.3389/fpsyg.2014.00691
- Quian, T., Jaeger, T.F., & Aslin, R.N. (2012). Learning to represent a multi-context environment: more than detecting changes. *Frontiers in Psychology, 3*(228), doi: 10.3389/fpsyg.2012.00228
- Reber AS. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*, 855–863.
- Regier, T. (2005). The emergence of words: attentional learning in form and meaning. *Cognitive Science, 29*, 819-865.
- Reisenauer, R., Smith, K., & Blythe, R.A. (2013). Stochastic dynamics of lexicon learning in an uncertain and nonuniform world. *Physical Review Letters, 110*(258701), 1-5.
- Romberg, A.R. & Yu, C. (2014). Interactions between statistical aggregation and hypothesis testing during word learning. In P. Bello, M. Guarini, M. McShane, &

- B. Scassellati (Eds.). *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1311-1316). Québec City, Canada: Cognitive Science Society
- Saffran, J.R., & Thiessen, E.D. (2007). Domain-general learning capacities. In E. Hoff & M. Shatz (Eds.). *Handbook of Language Development* (pp. 68-86). Cambridge: Blackwell.
- Sahni, S.D., Seidenberg, M.S., & Saffran, J.R. (2010). Connecting cues: Overlapping regularities support cue discovery in infancy. *Child Development, 81*(3), 727-736.
- Schmidt, S. R. (1991). Can we have a distinctive theory of memory? *Memory & Cognition, 19*, 523-542.
- Scott, R.M. & Fisher, C. (2012). 2.5-year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition, 122*, 163-180.
- Siskind, J.M. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition, 61*, 39-91.
- Smith, L.B., Suanda, S., & Yu, C. (2014) The unrealized promise of infant statistical word-referent learning. *Trends in Cognitive Sciences, 18*(5), 251-258.
- Smith, K., Smith, A.D.M., & Blythe, R.A. (2011). Cross-situational learning: An experimental study of word-learning mechanisms. *Cognitive Science, 35*, 480-498.
- Smith, L. B. & Yu, C. (2008) Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition, 106*(3), 1558-1568.

- Suanda, S.H., Mugwanya, N., & Namy, L.L. (2014). Cross-situational statistical word learning in young children. *Journal of Experimental Child Psychology*, *126*, 395-411.
- Suanda, S.H., & Namy, L.L. (2012). Detailed behavioral analysis as a window into cross-situational word learning. *Cognitive Science*, *36*(3), 545-559.
- Swingle, D., & Aslin, R. (2007). Lexical competition in young children's word learning. *Cognitive Psychology*, *54*(2), 99-132.
- Trueswell, J.C., Medina, T.N., Hafri, A., & Gleitman, L.R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology*, *66*, 126-156.
- Turk-Browne, N.B., Jungé, J.A., & Scholl, B.J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General*, *134*(4), 552-564.
- Turk-Browne, N.B., Scholl, B.J., Chun, M.M., & Johnson, M.K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience*, *21*(10), 1934-1945.
- Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *The Journal of Neuroscience*, *30*(33), 11177-11187.
- Valian, V. (2015). Bilingualism and cognition. *Bilingualism, Language, and Cognition*, *18*(1), 3-24.

- Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition*, *127*, 375-382. doi: 10.1016/j.cognition.2013.02.015
- Vlach, H.A. & Sandhofer, C.M. (2014). Retrieval dynamics and retention in cross-situational statistical word learning. *Cognitive Science*, *38*, 757-774.
- Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition*, *107*, 729-742.
- Vouloumanos, A., & Werker, J. (2009). Infants' learning of novel words in a stochastic environment. *Developmental Psychology*, *45*(6), 1611-1617.
- Wallace, W.P. (1965). Review of the historical, empirical, and theoretical status of the von Restorff phenomenon. *Psychological Bulletin*, *63*(6), 410-424.
- Wang, T., & Saffran, J.R. (2014). Statistical learning of a tonal language: The influence of bilingualism and previous linguistic experience. *Frontiers in Psychology*, *5*(953), doi: 10.3389/fpsyg.2014.00953.
- Yoshida, H., Tran, D.N., Benitez, V., & Kuwabara, M. (2011). Inhibition and adjective learning in bilingual and monolingual children. *Frontiers in Developmental Psychology*, *2*, 1-14. doi: 10.3389/fpsyg.2011.00210.
- Yu, C. & Ballard, D.H. (2007). A unified model of early word learning: Integrating statistical and social cues. *Neurocomputing*, *70*(13-15), 2149-2165.
- Yu, C., & Smith, L.B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414-420.
- Yu, C. & Smith, L. B. (2012) Modeling cross-situational word-referent learning: Prior questions. *Psychological Review*, *119*(1), 21-39.

Yurovsky, D., & Frank, M.C. (2015). An integrative account of constraints on cross-situational learning. *Cognition*, *145*, 53-62.

Yurovsky, D., Smith, L. B. & Yu, C. (2013) Statistical word learning at scale: The baby's view is better. *Developmental Science*, *16*, 959-966.

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