



ELSEVIER

Contents lists available at ScienceDirect

Cognition

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

Original Articles

Fast mapping word meanings across trials: Young children forget all but their first guess

Athulya Aravind^{a,*}, Jill de Villiers^{b,*}, Amy Pace^c, Hannah Valentine^c, Roberta Golinkoff^d, Kathy Hirsh-Pasek^e, Aquiles Iglesias^f, Mary Sweig Wilson^a Department of Linguistics and Philosophy, MIT, United States^b Department of Psychology, Smith College, United States^c Department of Speech and Hearing Sciences, University of Washington, United States^d School of Education, University of Delaware, United States^e Department of Psychology, Temple University, United States^f Department of Communication Sciences and Disorders, University of Delaware, United States

ARTICLE INFO

Keywords:

Word learning
Cross-situational learning
Propose-but-verify
Fast mapping

ABSTRACT

Do children learn a new word by tracking co-occurrences between words and referents across multiple instances (“cross-situational learning” models), or is word-learning a “one-track” process, where learners maintain a single hypothesis about the possible referent, which may be verified or falsified in future occurrences (“propose-but-verify” models)? Using a novel word-learning task, we ask which learning procedure is utilized by preschool-aged children. We report on findings from three studies comparing the word-learning strategies across different populations of child learners: monolingual English learners, Spanish - English dual language learners, and learners at risk for language-delay. In all three studies, we ask what, if anything, is retained from prior exposures and whether the amount of information retained changes as children get older. The ability to make a good initial hypothesis was a function of various factors, including language ability and experience, but across-the-board, children were no better than chance after a wrong initial hypothesis. This suggests that children do not retain multiple meaning hypotheses across learning instances, lending support to the propose-but-verify models.

1. Introduction

A fundamental question in developmental research concerns how knowledge changes over time. Nowhere is this question more pertinent than in the domain of word learning: children go from having effectively no recognizable words in their productive vocabulary before the age of 1 year to having a few hundred just a year or so later. The remarkable ease and efficiency with which children learn the mappings between phonological forms and their meanings might mislead one into thinking that word-learning is a simple task. But as famously pointed out by Quine (1960), even in the simplest cases—involving words denoting concrete nouns and the objects they represent, which most researchers agree can be learned observationally—the task is highly non-trivial. Any naturalistic learning situation has the potential for infinite referential ambiguity in word-to-meaning mapping. Much research in early child language has tried to demonstrate that the situation is less extreme, for instance, children may be predisposed to favor certain mappings over others, thus imposing constraints on the possibilities

(Golinkoff, Hirsh-Pasek, Bailey, & Wenger, 1992; Hollich, Golinkoff, & Hirsh-Pasek, 2007; Markman, 1990). Nevertheless, uncertainty is not eliminated altogether, and the question of how children learn the meaning of words given this state of affairs remains a core one in developmental research.

One response to the referential ambiguity problem is to say that the problem is only intractable if the learning situation is limited to one instance in which the word is uttered (Siskind, 1996; Smith & Yu, 2008; Yu & Smith, 2007 and many others). To illustrate the point, consider a child who hears the word ‘rabbit’ in a situation where a rabbit, a cat and a teddybear are present. Let us also assume that the myriad other possibilities imagined by Quine—the shadows, the lighting, the furri-ness, the ear length—are taken out of consideration by the constraints the child brings to the situation, such as the “whole-object” constraint (Markman, 1989). From this situation alone, it may be still be unclear what ‘rabbit’ means. But if the child encounters multiple situations where ‘rabbit’ is heard, and a rabbit is the only constant among the alternatives, then she may be able to identify the right meaning for the

* Corresponding authors at: Department of Linguistics & Philosophy, 77 Massachusetts Ave, 32D-808, Cambridge, MA 02139, United States (A. Aravind). Department of Psychology, 110 Elm Street, Bass Hall, 401, Northampton, MA 01063, United States (J. de Villiers).

E-mail addresses: aaravind@mit.edu (A. Aravind), jdevilli@smith.edu (J. de Villiers).

<https://doi.org/10.1016/j.cognition.2018.04.008>

Received 8 October 2017; Received in revised form 8 April 2018; Accepted 9 April 2018

Available online 25 April 2018

0010-0277/ © 2018 Elsevier B.V. All rights reserved.

word. This is the core idea behind *cross-situational word learning* (CSM) models: if the learner can track statistical co-occurrence rates across many situations, she is bound to gradually converge on the target mapping.

To successfully use this learning mechanism, the learner must store and compare multiple parallel hypotheses across several learning situations in order to accumulate enough evidence eventually to select the best hypothesis. But can human learners do this? This issue has been investigated experimentally with both adult and child populations. In a typical task, participants have to learn the meaning of several new words in ambiguous situations. Though each individual trial is uninformative, adult participants selected the correct referent at above-chance levels by the end of the study session (McMurray, Horst, & Samuelson, 2012; Vouloumanos, 2008; Yurovsky, Fricker, Yu, & Smith, 2014). The paradigm has also been adapted to test whether children use the same associative learning strategies as adults seemed to be using. Smith and Yu (2008) presented 12- to 14-month-olds with two novel words and two novel referents in an ambiguous training trial. In a later test trial, children saw one of the two referents along with a novel referent. Children preferentially looked at the target in the test trials, despite both the training and test trials being individually ambiguous. The difficulty arises in interpreting this finding as suggesting that the learners were indeed keeping track of information about multiple candidate referents for the same word. There are different possibilities for how experience can accumulate.

An alternative account to the cross-situational learning model posits that word-learning is a “one-track” process, where learners maintain a single hypothesis about the possible referent for any given novel word, which may be verified or falsified in future occurrences. Note that the only evidence that the guess is false is that the hypothesized referent does not appear on the next trial. In the event of falsification, a new candidate referent is hypothesized. In the kind of experiment described, this too would lead to increased success by the final trial, on average. *Propose-but-verify* (PBV) models of word learning posit that nothing is stored from one trial except the learner’s latest guess. Supporting evidence comes in part from experiments with adults showing that participants maintain their hypotheses over time if they made an initial correct guess, but pick a new referent at chance among the available candidates if their initial guess was incorrect (Medina, Snedeker, Trueswell, & Gleitman, 2011; Trueswell, Medina, Hafri, & Gleitman, 2013). In other words, participants do not seem to retain any information about alternative targets from past trials. Only one study, by Woodard, Gleitman, and Trueswell (2016), has thus far tested predictions of propose-but-verify with children, and their results also support the model. In the study, 2- and 3-year old children were shown an initial trial involving two novel animals accompanied by a novel word and asked to select the referent for the word. In a second exposure, they were shown a new pair, one of which was either (i) the same animal they selected (“Same” condition) or (ii) the animal they had failed to select previously (“Switch” condition) (Fig. 1). Whereas children could remember their earlier hypothesis in a subsequent trial in the Same condition, they showed no indication of recalling the unselected alternative in the Switch condition. However, as Fig. 1 shows, the other choice was an unknown object that had been seen before, against a familiar alternative (Trial 1). Woodard et al. had designed the task this way to ensure that any above-chance responding observed would not be due to a simple familiarity effect. Nevertheless, in avoiding one confound, another one could have been introduced: the child on the test trial could have recalled both unselected alternatives, but lost track of the situation in which each was presented.

More recently, hybrid accounts have been proposed on which learners may strongly favor a single hypothesis, but also extract some information about alternative candidates (Yurovsky & Frank, 2015; Roembke & McMurray, 2016; Stevens, Gleitman, Trueswell, & Yang, 2017). Yurovsky and Frank’s (2015) experimental results with adults were consistent with an integrative model, where learners allocated a

fixed amount of attention to a single hypothesis but distributed the rest evenly among the remaining alternatives. The amount of attention allocated to each alternative candidate decreases as the complexity of the scene (i.e. number of potential candidates) increases, so whereas performance after failure was above chance in low-complexity trials, it did not differ from chance in high-complexity trials. There is further experimental evidence that adult participants consider information about previously seen alternatives and even extraneous information from previous exposures.¹ In an eye-tracking study by Roembke and McMurray (2016), participants looked more at previously seen competitors even when they would eventually choose the target. Dautriche and Chemla (2014) found that adults also retained additional, irrelevant information e.g. about the context in which the stimuli appeared and the spatial location of the target referent. The “Pursuit” model presented in Stevens et al. (2017) takes a more nuanced angle on Propose-but-verify. The critical way in which Pursuit differs from Propose-but-verify is in that on Pursuit, learners retain the disconfirmed hypotheses over the course of many learning instances and ultimately choose the best one, i.e., the one that was confirmed most often.

Overall, these hybrid accounts view word learning as a complex, multidimensional process that incorporates aspects of both associative, cross-situational word learning and Propose-but-verify. However, it is an open empirical question whether children’s word learning involves both “one-trial” learning as well as associative mechanisms. The only previous study with children is the one we discussed above, and as we pointed out, we cannot be fully certain that the alternatives were not in fact remembered. The rest of the work target adult populations, but there are reasons to doubt the validity of inferences from findings in adult work to children: Adults and children differ substantially in general cognitive abilities such as memory and attention, and children may be less able to remember alternative hypotheses or other information from the initial exposure.

The present study uses a novel word-learning task to contribute to the question of how children learn new words in more complex settings. The task was not designed as an experiment, but rather, it was developed as a test of *fast mapping* ability (Carey & Bartlett, 1978) for a language screening assessment, QUILS (Golinkoff, de Villiers, Hirsh-Pasek, Iglesias & Wilson, 2017). The design is interestingly different from previous studies of novel word learning. The number of candidate-choices is greater, but a verbal prompt, providing semantic and contextual linguistic cues, was included to aid target selection. In both respects, the task more closely simulates a child’s real-world experience of learning novel words and offers a more ecologically valid test of their learning mechanisms than the usual experimental set-up. Furthermore, on the test trial, all except the target represented objects the child had not seen previously. In Woodard et al., the alternatives had been presented before to control for familiarity.

We report on results from three large-scale studies that allow us to compare the various learning models across different populations of child learners. In all three studies, we ask what, if anything, is retained from prior exposures and whether the amount of information retained changes as children get older. The answers to these questions will help us arbitrate among cross-situational learning, propose-but-verify, and more hybrid, models of learning. We also ask whether the particular learning strategies employed vary across different types of learners. We compare the behavior of typically developing monolingual learners of English to typically developing dual language learners, as well as atypical learners to ascertain to what extent the core underlying learning mechanisms are shared across these different populations.

Study 1 presents results from typically developing monolingual

¹ Trueswell et al. (2013) criticize these findings on the grounds that the tasks themselves are highly unnatural: for instance, the set of stimuli are often “closed”, revolving around a small set of objects participants see repeatedly over the course of the study. Consequently, they argue that although adults *can* in principle extract more than one meaning, they may only do it under very contrived experimental settings.

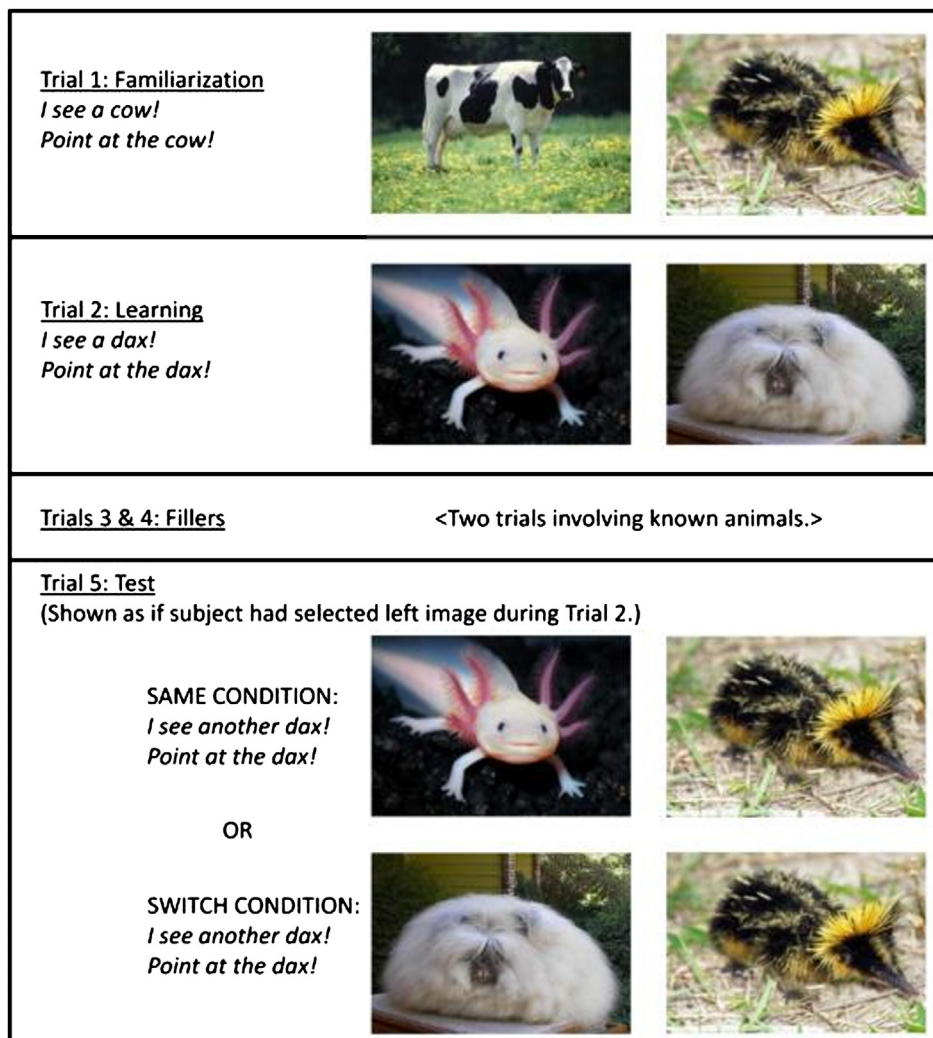


Fig. 1. Example Trial, Woodard, Gleitman, and Trueswell (2016).

learners of English in a large sample of 659 children. So far, most research on models of word learning has been conducted with monolingual learners who were typically developing. In the process of validating the two screeners we have the opportunity to see if there are significant differences in the process from other groups acquiring English. In Study 2, we compare these results to a large group (N = 568) of dual language learners of English and Spanish, who were administered the task in both their native languages as part of the development of the QUILS-ES, for dual language learners (Iglesias, de Villiers, Golinkoff, Hirsh-Pasek & Wilson, in press). We discuss why the situation might be different for dual language learners, who may have to keep more hypotheses about the labels for things in mind. Finally, Study 3 investigates the word-learning strategies employed by atypical language learners, and ask if and how the process differs. Potential differences in the learning mechanisms between typical and atypical learners address questions of general import, such as how much extant knowledge and abilities impact the formation of strategies to learning (see e.g. Snedeker & Gleitman, 2004).

2. Study 1

Study 1 presents the results from a large sample of monolingual English-speaking children presented with a novel word-learning task. The large and varied sample allows a thorough exploration of the question of how children learn new words in more complex settings, using a combination of linguistic and contextual clues to narrow down alternatives. The study allows for a direct comparison of the cross-situational and propose-but-verify learning models. By tracing an individual's choices across trials we were able to ask if the child could remember only her first hypothesis, or hold onto alternatives and recover from mistakes. Some of the previous work predicts changes over the lifespan in the ability to keep prior or rejected hypotheses in mind, as a function of executive function, attention, or memory (Yurovsky & Frank, 2015), as well as social information processing (Yurovsky & Frank, 2017). Given that these abilities continuously improve between infancy and 5 years of age, an important secondary question we will explore is whether children improve their guesses with age.

Table 1
Age groups, Study 1.

	Frequency	Percent
3,0–3,11 (3yr-old-group)	213	31.6
4,0–4,11 (4yr-old-group)	315	46.7
5,0–5,11 (5yr-old-group)	146	21.7

Table 2
Gender, Study 1.

	Frequency	Percent
Male	322	47.8
Female	352	52.2

Table 3
Subtests included in QUILS.

Vocabulary product	Grammar product	Process
Known nouns	wh-questions	Fast mapping nouns
Known verbs	Past Auxiliary/copula	Fast mapping adjectives
Prepositions	Prepositional phrases	Converting active to passive
Conjunctions	Embedded complements	Syntactic bootstrapping

2.1. Participants

Participants were 674 monolingual children, aged 3;0 to 5;11, balanced by gender. Age and gender information are summarized in Tables 1 and 2. The children came from various parts of US, and were racially and ethnically diverse.

2.2. Materials and procedure

The novel word-learning task reported in this study was extracted from a language assessment tool, the Quick Interactive Language Screener (QUILS/QUILS-ES; Brookes Publishing, 2017). The QUILS uses touchscreen technology to evaluate preschool-aged children’s receptive language skills. The overall structure of QUILS is schematized in Table 3.

The assessment included measures of both “products” of knowledge, i.e. knowledge about words and syntax that children acquire through experience with their native language, and of learning “processes,” their ability to learn new words and grammatical structures with limited exposure. The data reported in this paper, representing the “Fast Mapping Nouns” subtest of QUILS, were obtained as part of the data collection conducted for the development of this tool.

Participating children received QUILS in one or two brief sessions, one of which included 5 (successive) items of Fast Mapping Nouns. Children’s responses to the touchscreen were automatically recorded. Each Fast Mapping item was associated with two trials. In Trial 1, the “noun-learning” trial, participants are introduced to a novel object embedded in an array and hear a novel noun. In Trial 2, the test trial, they were expected to identify another referent for the newly learned noun. The test trial immediately followed the noun-learning trial. Because preschool children tend to already be quite adept at mutual exclusivity tasks, the selection set in the noun-learning trial included two novel objects, one of which was the target, and two familiar objects (Fig. 2). Choice of target was guided by a linguistic cue, which isolated just one of the novel objects. The foils or competitors in the learning trial included (i) one novel object, which did not satisfy the property in the linguistic cue (e.g. blueness), and (ii) two known objects, both of

The *fep* is blue! Can you find the *fep*?

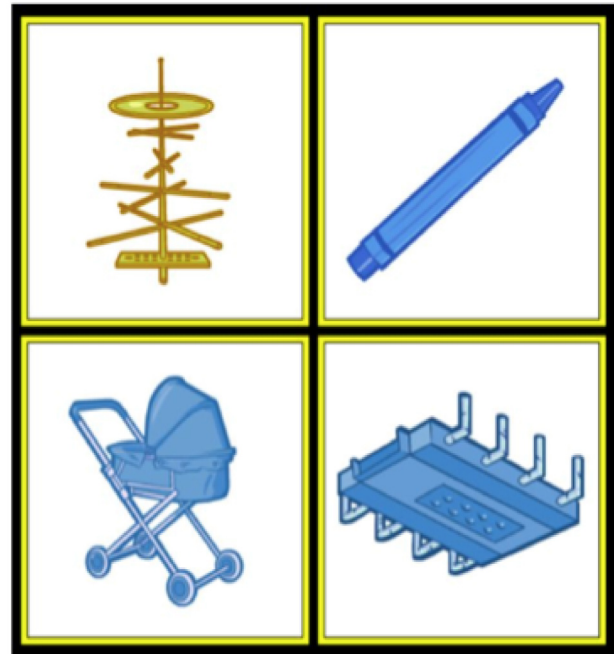


Fig. 2. Trial 1, “Fep”.

Can you show me another *fep*?



Fig. 3. Trial 2, “Fep”.

which possessed the relevant property. Thus, the child was required to integrate mutual exclusivity and linguistic knowledge to converge on the correct object. Moreover, unlike much of the past word-learning tasks in CSM and PBV literature, the initial trial was not ambiguous.

In Trial 2, the test trial, the linguistic cues were removed, requiring children to use only their newly acquired knowledge of the noun to a new context. Thus, continuing the example from Fig. 2, children would be shown the array in Fig. 3 and asked, “Can you show me another

feh?”. Note that only the target is repeated in the test trial, but not exactly. Changes included color, orientation and size. In this aspect, too, our design was different from previous studies, where there were at least two intervening, irrelevant items between learning and test trials. Here, on the other hand, Trial 2 occurs immediately after the learning trial, giving children the best chance at remembering things from one trial to the next.

The design of this study is suited to adjudicate between CSM models and PBV models, as well as hybrid models, as it allows us to explore whether *any* information, either about the alternative choices or the linguistic and contextual cues, are retained after initial exposure. More specifically, if the child failed on Trial 1, do they have a higher than chance likelihood of retaining the information about the target object, which is the only thing in common from the first trial among the four Trial 2 foils? CSM posits that children carry knowledge of past potential targets into each word-learning exposure, i.e. they will use information from previous trials to inform their performance on secondary trials. On the other hand, PBV predicts that if children have made an incorrect conjecture, their behavior on Trial 2 will not be better than chance. Additionally, does children’s ability to make good guesses and retain information vary as a function of age and general language ability?

2.3. Results

2.3.1. Overall results

Recall that our design differed from previous studies in that there was a correct choice, even at first exposure. Moreover, since converging on the target implicates the ability to integrate both linguistic cues and perceptual/contextual information regarding the array, we expected children’s performance on this task to improve with age. Fig. 4 displays the performance at each Age group on the initial trial, i.e. how often do they use the relevant cues to pick out the correct object initially? Fig. 5 depicts how overall ability to choose the correct object on Trial 1 and retain and extend it on Trial 2 differs by Age group. In both cases, we see improvement with age: whereas the 3 year olds are not so good at the task, children at age 5 are doing considerably better (though even 5-yr-olds are not at ceiling).

2.3.2. Accuracy on Trial 2 as a function of Trial 1

CSM and PBV models make different predictions about retention

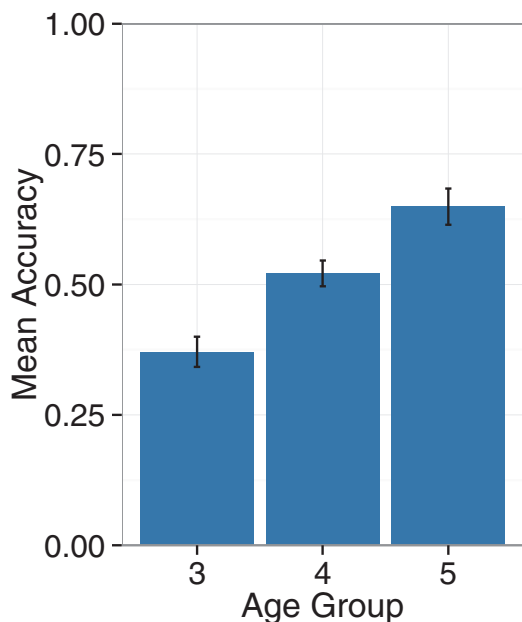


Fig. 4. Accuracy on initial trial.

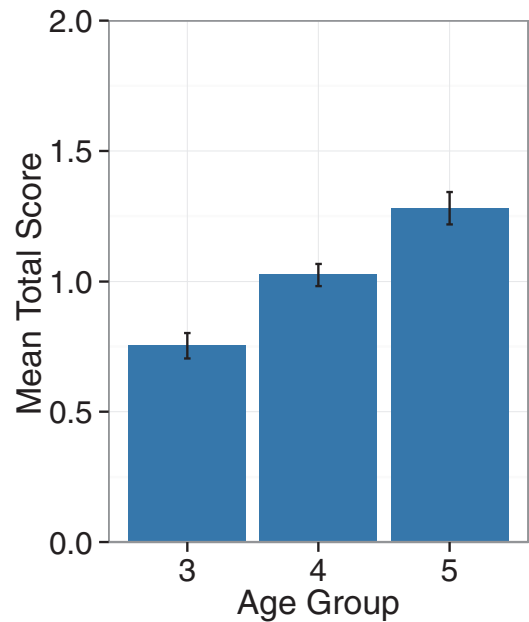


Fig. 5. Growth with age on total score.

across trials, so we now turn to the question of whether children maintain their hypotheses across trials. Fig. 6 represents mean accuracy on Trial 2 as a function of success (1) or failure (0) on Trial 1, split by Age Group (3-yr-olds, 4-yr-olds, and 5-yr-olds). It is clear that children who succeed on Trial 1 are successful on Trial 2, but crucially, the performance of children who fail on Trial 1 hovers around chance (25%) on Trial 2. Observe also that whereas the ability to choose the target again after an initial success seems to steadily increase with age, there is no improvement with age on Trial 2 for those who were unsuccessful on Trial 1.

To evaluate these trends statistically, we first fit a mixed-effects logistic regression predicting Trial 2 accuracy as a function of Trial 1 accuracy. We also included as interacting fixed factors Age Group (helmert-coded) and overall language ability, as measured by the

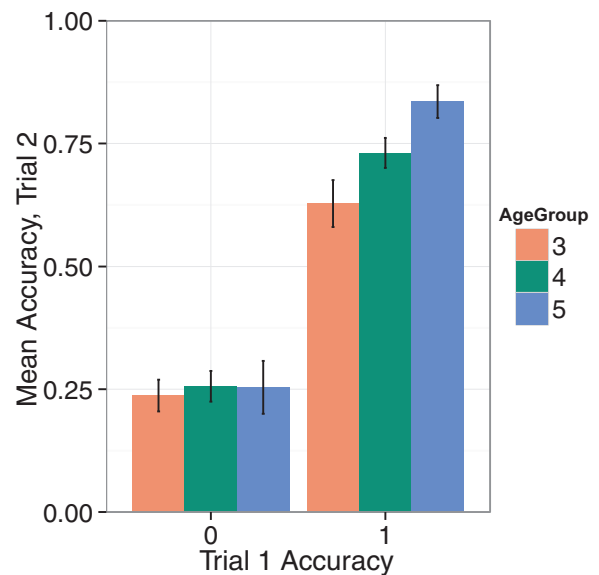


Fig. 6. Accuracy on Trial 2 by Trial 1 success/failure.

child’s average score on the Vocabulary and Grammar Product modules of the QUILS (so as not to include Fast Mapping Nouns itself). Finally, we included random intercepts for participant and item. Results reveal a significant three-way interaction of Trial 1, Age Group and Language Ability ($\beta = -0.119$, $SE = 0.040$, $z = -3.003$, $p = 0.003$), such that age and language ability make a difference only if Trial 1 was a success and moreover, language ability makes a greater difference at 3- and 4-years of age compared to 5- years of age.

In addition to these overall effects, we were also interested in whether performance on Trial 2 was different from chance for Trial 1 passers vs. failers. We examined the two groups separately, and applied two tests. First, we used a one-sample Wilcoxon test to compare mean accuracy on Trial 2 to a null hypothesis of chance-level performance (25%). Participants who succeeded on Trial 1 succeeded on Trial 2 at a rate of 73.6% ($SE = 1.1\%$), which was significantly greater than chance ($V = 107032$, $p < 0.001$). Participants who failed on Trial 1 succeeded on Trial 2 at a rate of 24.8% ($SE = 1.0\%$), which was not significantly different from chance ($V = 613824$, $p = 0.99$). Of course, these statistical tests do not account for variance due to specific items. Therefore, we conducted a second pair of analyses using a logistic mixed regression that accounts for such variation. Our models, run separately for the subgroups of Trial-1 passers and failers, included the intercept as a fixed effect, along with random intercepts for participant and item. We then compared whether the estimated intercept, after by-subject and by-item random variation was accounted for, was different from chance. The results of the analysis of Trial 1 passers showed the intercept to be estimated at 1.2657 ($SE = 0.11$), corresponding to a raw value of 78% success rate with 95% CIs of [1.057, 1.474], corresponding to raw values of [74.2%, 81.4%]. This finding shows that overall performance on Trial 2 was indeed better than chance for those who succeeded on Trial 1. For Trial 1 failers, the estimated intercept was -0.8907 ($SE = 0.22$), corresponding to a raw value of 24% success rate. 95% CIs were $[-1.258093 - 0.9650822]$, corresponding to raw values of [22%, 28%]. Chance level, 25% (-1.099 after logit transformation), does fall inside the confidence intervals, suggesting that the success rates for Trial 1 failers was not reliably different from chance.

2.3.3. Nature of Trial 1 errors

An important critique remains: could the random results on Trial 2 following an error on Trial 1 reflect a problem with the participant who makes such an error on Trial 1? There are two ways this could play out: the first is that it is a subset of children who are contributing to the error analyses, namely, children who are shy, uncertain, or simply unable to carry out the task and guess randomly. The second is that the pattern of error on Trial 2 is not a product of only certain children, but of any child who resorts to a bad choice on Trial 1. In other words, the Trial 2 choices may be contaminated by all kinds of other factors that could make performance worse (e.g. inattentiveness) that are not relevant to the theoretical question of interest. The fact that our task is different from others in this literature—there is in fact a right answer on Trial 1—makes this question even more urgent. In Woodard et al., or in Yurovsky and Frank, the children’s first guess is called a “hypothesis”, and there is no way to distinguish randomness from deliberate choice. However in the present study it is obvious when a child chooses wrongly on Trial 1. When children get that trial wrong, do they continue in the same vein, namely with errors or confusion? That is, children could be behaving strangely on Trial 2 in a way that would not be predicted by either cross-situational learning or propose-but-verify: their responses could simply indicate either that they were inattentive on Trial 1, and as a result, miss out on crucial information that could help them carry on in Trial 2.

We attempt to address this issue in two steps. We first ask whether

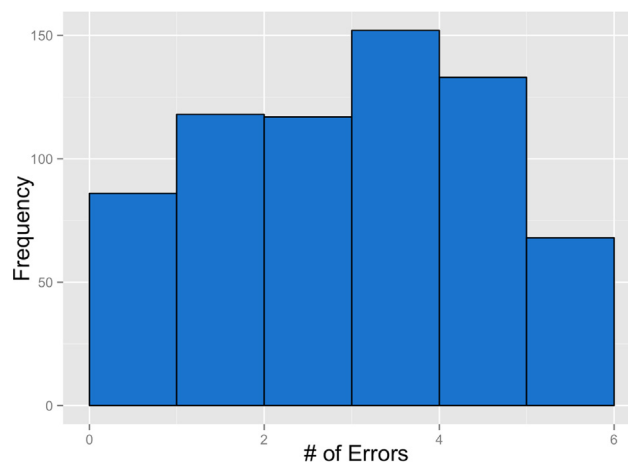


Fig. 7. Distribution of errors on Trial 1 across the sample.

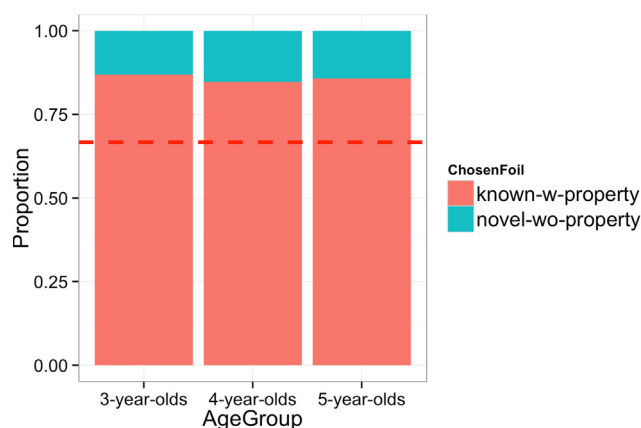


Fig. 8. Trial 1 foil choices by age group.

the effect is carried by a group of children who are systematic failers. If the trends we observe are due to general distracted/shyness of the child, we might see an all-or-nothing pattern, with some participants systematically getting things wrong (due to e.g. shyness) and others systematically getting things right. This is not the case—Fig. 7 demonstrates that errors on Trial 1 are normally distributed, with most children making a couple of errors.

Almost all children, then, are contributing to the error results. But what is the error pattern on Trial 1? On Trial 1, if the child failed to choose the target, she had two opportunities (66% chance) to choose a known object with the relevant property described by the linguistic cue, and one opportunity to choose a novel object without the relevant property (33% chance). However, children show a “known-w/property” bias, a bias that is stable across age group (Fig. 8). This suggests that (i) children’s first trial choices, even when they are getting it wrong, are principled, in that they use the linguistic clue and (ii) mutual exclusivity is a more violable bias at even the youngest ages, compared to the linguistic clue. This pattern would not be expected if children were failing to pay attention on Trial 1 and making a random guess. Therefore the first choice is not random.

Continuing in this vein, we asked whether the error patterns might be distinct for those children who made more errors on Trial 1. We divided those children who made errors at all into two groups based on Trial 1:

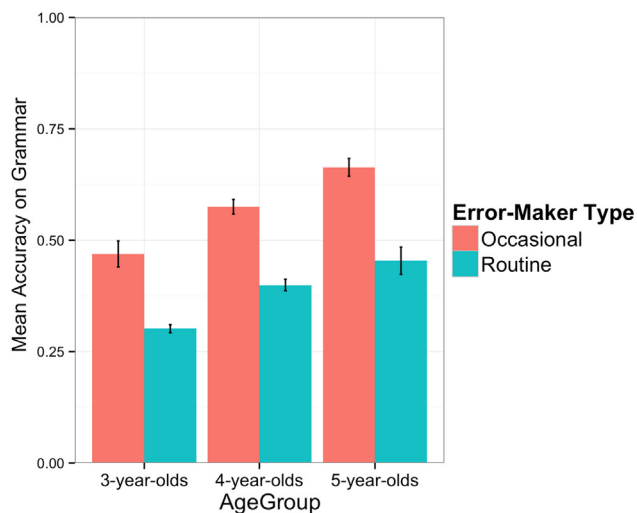


Fig. 9. Grammar ability by age and type of error maker.

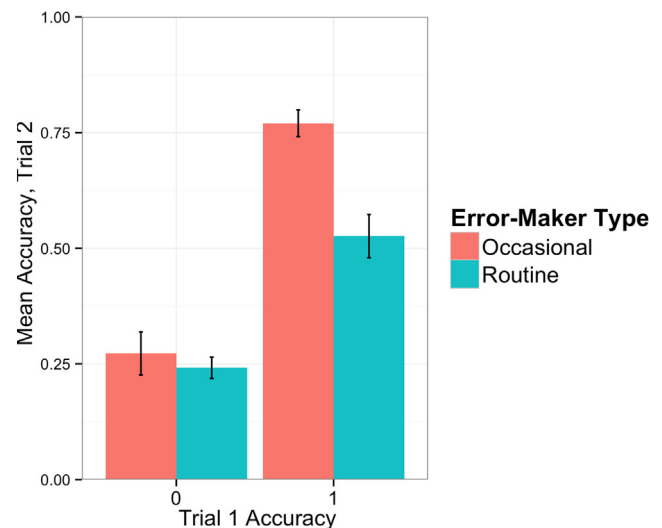


Fig. 10. Accuracy on Trial 2 by Trial 1 success/failure.

- (i) those that made 3 or more errors (the “Routine-Error-Makers”)
- (ii) those that made only 1 or 2 errors (the “Occasional-Error-Makers”)

Children belonging to these two groups have different profiles when it comes to grammatical ability and age. We conducted a mixed-effects logistic regression that examined the likelihood of falling into the “Routine-Error-Maker” category as a function of Age Group and Grammatical Ability. We found that children were less likely to fall into this category at older ages ($\beta = -8.45$, $SE = 1.38$, $z = -6.123$, $p < 0.001$). Moreover, Grammatical Ability (as measured by the grammar module of QUILS) also strongly negatively predicted the likelihood of being a Routine Error-Maker ($\beta = -8.93$ $SE = 3.26$ $z = -2.73$ $p = 0.006$). There was no reliable interaction of Age and Grammatical Ability. In sum, then, the older the child and the more grammatically capable, the less likely she is to fall under the “Routine-Error-Maker” category (See Fig. 9).

Having established that the two groups of error-makers differ in fundamental ways, we then asked if and how the two error groups varied in their responses on Trial 2. As shown in Fig. 10, the two groups differed in how successful they were at extending an initial good hypothesis to a second situation: routine error makers were markedly less successful at choosing the target on a second trial even after a successful first trial. This is perhaps not surprising, given that they tend to be younger and less linguistically sophisticated. Crucially, however, the two groups do not differ after a failed first trial. Both groups are no different from chance on Trial 2 following an error on Trial 1.

These error analyses demonstrate that the failure to maintain alternatives from Trial 1 – and therefore no savings across contexts – is not an artifact of certain inattentive or foolish participants, nor are children’s choices on Trial 1 random. Regardless of whether children make a few or a lot of the wrong choices on Trial 1, their performance is equally at chance on Trial 2 following such a choice.

2.4. Discussion

We started with the question of what, if anything, children remember from one trial to the next. The main observation from our results is that whereas children successfully remember their *correct* guess from Trial 1 on Trial 2, they do not seem to be remembering anything else. Even though other linguistic cues guided the search, if the child gets that wrong, they do not recover on Trial 2 by recognizing

a missed alternative, even though Trial 2 immediately followed the learning trial. Moreover, whereas children improve in their ability to remember their initial correct hypothesis, there is no evidence that children improve between 3 and 5 years of age in having cross-trial savings about alternative candidates. This is despite the fact that children in this study, unlike in Woodard et al., could in fact have used a “familiarity strategy” to achieve above-chance performance, because only the target was repeated. The fact that they did not means that is a further strong argument for the Propose but Verify model. For the “familiarity strategy” to take effect, the candidate has to not just be familiar, but relevantly so; the alternative candidates were not relevant for the goal at hand, e.g. identification of the *lep*, and are simply not remembered.

3. Study 2

Study 2 asks the same questions about children of the same age simultaneously learning two languages, English and Spanish. The process of fast mapping has been very little investigated in dual language learners, yet the process of acquiring words for objects has been acknowledged to be more complicated. In a monolingual learner, it has been proposed that there is an initial bias towards mutual exclusivity, namely, one word per object. That is precisely what allows the child in our task to discard the known object with the correct property as the target. However, as our error analyses reveal, by age 3, mutual exclusivity is better characterized as a soft bias: error-makers are happy to overlook it in a manner that contrasts with their reliance on the linguistic cue. This pattern is unsurprising: as the child’s vocabulary is enriched, she must come to learn that a single referent may be described by a series of overlapping, hierarchically related labels (e.g. *Fido*, *beagle*, *dog* and *animal* for the family dog). Thus, at a certain stage in development, children must not only be able to use the mutually exclusivity assumption to their benefit, but also override it as needed. In the case of a dual language learner, the need to override the mutual exclusivity assumption is more urgent. Two different words might often be heard in the presence of an object, one in each language, and each must be stored as a hypothesis awaiting further confirmation. Not only that, but the word from one language must not be used to disconfirm the word in the other language that was heard earlier. *Perro* cannot supersede *dog* or vice versa: both must be kept. The multiple mappings

make the situation much more complex for dual language learners, especially for those too young to even be cognizant of the existence of two separate linguistic systems in their environment (e.g. Kalashnikova, Mattock, & Monaghan, 2014). To make things worse, “false friends” exist, where the same phonological form maps to different meanings in each language, for example the word *éxito*, which is unrelated to the English *exit* and means ‘success’ in Spanish. It is critical, therefore, that word-learning research accommodates the case of bilingual input in which assumptions of mutual exclusivity do not obtain (Byers-Heinlein & Werker, 2009; Grosjean, 2008, 2010; Kalashnikova et al., 2014; Marian & Shook, 2012).

There is some recent research suggesting despite these potential challenges, dual language learners behave rationally in fast mapping environments. For example, Kandhadai, Hall, and Werker (2017) asked whether young dual language learners are less likely to apply mutual exclusivity, or whether they entertain more hypotheses about a new word’s meaning, for example, that it might be a second category term for a familiar object. The task involved was eye tracking, so the stimuli and choices were quite constrained, and the interpretation of eye gaze behavior is uncertain. Nevertheless, presented with a familiar object (dog) in an unusual color (aqua), the 17–18-month-old monolinguals showed a tendency to assign the new word to the color. However dual language learning infants were more prone to treat the new word as a second label for the category DOG. The authors credit this difference to the language experience of the two groups.

Other research suggests that bilingualism in fact promotes novel word-learning abilities. For instance, Kaushanskaya and Marian (2009) tested Spanish-English and Mandarin-English bilinguals, along with monolingual controls. All groups were taught several novel words and were tested on recall and recognition measures immediately after learning and after a 1-week delay. Both bilingual groups outperformed monolinguals in all measures, leading the authors to conclude that there is a general bilingual advantage for novel word learning. A study by Poeysel and Weiss (2016) looks specifically at the issue of cross-situational word-learning in adult bilinguals, either Spanish-English or Chinese-English. Their first experiment looked at the success in learning novel word-object mappings given various distributions of the input, and found very little difference across the groups in their rates of success. However, the bilingual groups showed some superiority in recall to the monolingual group in the second study that used 2 to 1 mappings, that is, two words per object. The researchers conclude that the bilinguals showed greater flexibility in handling multiple mapping across time, compatible with a weakening of the constraint of mutual exclusivity. Though these studies involve adult participants and the focus is not on underlying mechanisms of word-learning per se, the question naturally arises: How well do dual language learners at the preschool age learn new words in each of their languages? Are they less able on Trial 1 to use mutual exclusivity to narrow their choices? And importantly, do dual language learners have a greater tendency to hold onto rival hypotheses than monolingual learners?

Study 2 takes advantage of data collected in the service of making a dual language version of the QUILS screener, the QUILS-ES (Iglesias, de Villiers, Golinkoff, Hirsh-Pasek & Wilson, in press). The same subset of fast Mapping Nouns was included for each of the different languages of the screener, which has a Spanish portion and an English portion. The items were chosen during piloting to be optimized for each particular language in terms of discriminability across age and ability level as total items (i.e. Trial 1 and Trial 2 combined). The final QUILS-ES thus has four items of fast mapping nouns in each language, with no overlap of items, for a total of eight. Four of these items, two in each language, are ones that also occur on QUILS. Hence Study 2 comes as close as possible to a matched study for Study 1.

3.1. Participants

The children in study 2 were dual language learners of English and Spanish, recruited from a variety of daycare sites for the norming of a new screener, QUILS-ES. Judged by schools and parents as dual language learners, the sample of 568 participants for the QUILS-ES is shown in Table 4 for age and Table 5 for gender.

Table 4
Age groups, Study 2.

	Frequency	Percent
3,0–3,11 (3yr-old-group)	148	26.1
4,0–4,11 (4yr-old-group)	232	40.8
5,0–5,11 (5yr-old-group)	185	32.6
Not reported	3	0.53

Table 5
Gender, Study 2.

	Frequency	Percent
Male	278	48.9
Female	285	50.1
Not reported	5	0.88

3.2. Results

Do dual language learners differ from monolingual learners in their ability to identify the target referent? Is the contingency of Trial 2 success on Trial 1 success different for dual language learners, and are there differences across the two languages being learned? Fig. 11 displays the rates at which dual language learners at the different ages converge on the right object in the initial trial. At the early ages, dual language learners are comparable to monolingual learners, but 5-year-olds who are learning two languages are choosing the right object at lower rates than their monolingual counterparts (see Fig. 4 above). This finding is in keeping with previous work showing that the reliance on

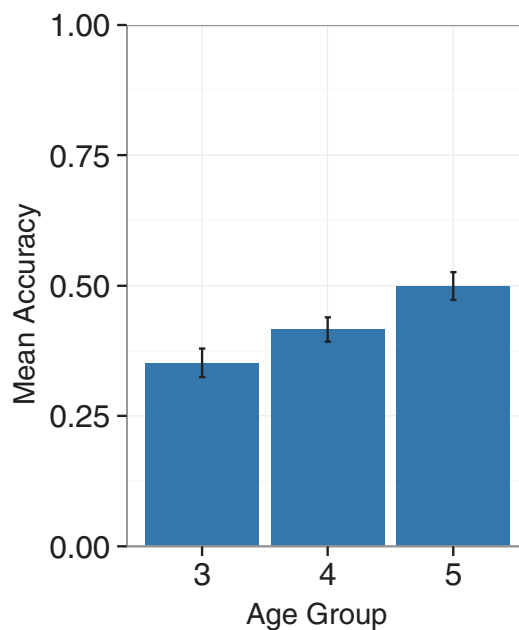


Fig. 11. Dual language learners’ performance on initial trial.

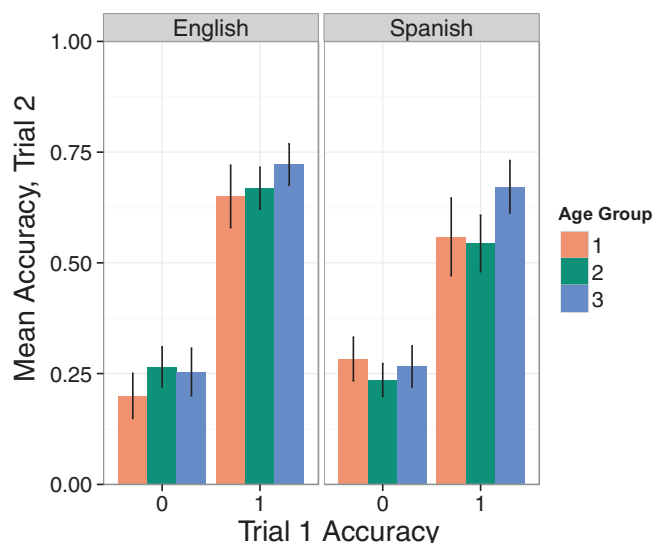


Fig. 12. Accuracy on Trial 2, dual language learners.

mutual exclusivity weakens in bilingual children as they become older and more experienced dual language learners (Kalashnikova et al., 2014).

Notice, however, that the patterns of Trial 2 success or failure as a function of Trial 1 are the same for this population, and there do not seem to be differences between the two languages (Fig. 12).

As in Study 1, we fit a logistic mixed-effects regression predicting Trial 2 accuracy as a function of Trial 1, as well as Age Group. We also included Language (Spanish/English) as a fixed factor for the dual language learners). We found a significant interaction of Trial 1 and Language ($b = 0.12$, $SE = 0.04$, $z = 2.823$, $p = 0.004$), such that accuracy on Trial 2 after a successful Trial 1 was greater in English than in Spanish. We also find an overall improvement in performance, for both languages, between Age Groups 4 and 5 ($b = 0.09$, $SE = 0.03$, $z = 2.728$, $p = 0.006$). However, between Age Groups 3 and 4, there was only improvement in English ($b = 0.11$, $SE = 0.05$, $z = 2.082$, $p = 0.03$).

3.3. Discussion

The performance of dual language learning children shows the same pattern in each of their languages on the same task, though with slightly weaker performance in Spanish. There is steady improvement across Age Group in getting Trial 1 correct on the basis of the cues provided, and also in extending that correct choice to the second Trial. However there is no evidence of cross-situational learning when the guess on Trial 1 was incorrect. The fact that dual language learners will, over time, need to establish two words in two languages for the same referent does affect their ability to identify the target referent initially, especially at later ages. However, it does not seem to impact what information they do and do not carry over to later learning instances; here, they perform just like the monolingual participants.

4. Study 3

Studies 1 and 2 demonstrated that typically developing monolingual and bilingual learners do not differ qualitatively in their word learning abilities. Children in both groups (and bilingual children in both their languages) showed better-than-chance performance on Trial 2 after a

successful Trial 1, but were no different from chance after a failure on Trial 1. We took these findings to indicate that both groups store their initial hypothesis about potential referents but no information about the alternative candidates. Our goal in Study 3 is to assess whether these patterns are stable across populations who can otherwise be shown to have different language-learning abilities.

Findings from one of the previous studies on cross-situational learning with children, Yu and Smith (2012), raise the possibility that linguistic skills might make a difference for what, and how much, information is stored from one trial to the next. Their study, which focused on word-learning abilities in 12–14 month olds, found that infants with the most advanced vocabularies for their age (as measured by the MCDI) were the ones who showed the highest rates of learning in a cross-situational environment. In a recent paper, Vlach and Debrock (2017) find a similar effect of overall language ability (as measured by the PPVT) on cross-situational word learning in preschoolers. In addition to the normative group tested on QUILS, we also tested a subgroup of children at risk for language delays, who have lower productive vocabularies and syntactic abilities than typical children in the same age brackets. In Study 3, we ask whether these children, who are below the typical range in their overall language ability, behave any differently than typically developing children in our word-learning task.

4.1. Participants, materials and procedure

The QUILS was given to 16 children identified as having Individual Education Plans, IEPs, because of language difficulties. All were receiving the services of SLPs in their schools or day care centers. There were 6 girls and 10 boys and their age and gender data are contained in Tables 6 and 7.

Table 6
Age Groups, Study 3, IEP group.

	Frequency	Percent
3,0–3,11 (3yr-old-group)	8	50
4,0–4,11 (4yr-old-group)	4	25
5,0–5,11 (5yr-old-group)	4	25

Table 7
Gender, Study 3, IEP group.

	Frequency	Percent
Male	10	62.5
Female	6	37.5

4.2. Results

Our findings are represented in Fig. 13. Though children with IEPs perform at lower rates overall than typically developing children, the trends are strikingly similar across the two groups: children who fail on Trial 1 hover around chance on Trial 2, whereas those who succeed on Trial 1 are above chance.

To assess the significance of these findings, we first fit a mixed-effects logistic regression, predicting Trial 2 Accuracy as a function of Trial 1 Accuracy and AgeGroup. No interaction terms were included, as the more specified model did not converge given the small sample. We found a significant main effect of Trial 1, with a successful Trial 1 leading to greater accuracy rates on Trial 2 ($b = 0.6320$, $SE = 0.267$, $z = 2.368$, $p = 0.0179$). We ran the same procedures to assess whether performance differed from chance for passers and failers. Wilcoxon tests on Trial-1 either group showed that: (i) Trial 1 passers were found to be above chance ($V = 437$, $p < 0.001$), but (ii) failers were not different from chance ($V = 581$, $p\text{-value} = 0.5322$). Intercept-only mixed effects

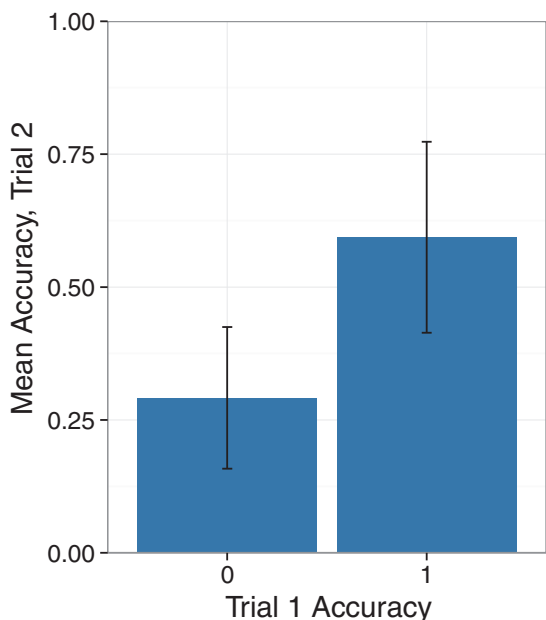


Fig. 13. Trial 2 success rates, children with IEPs.

models confirmed this: the estimated intercept was 0.3514, corresponding to a raw score of 58.6, with chance value falling outside of the 95% CIs, [−0.945, 1.648] (raw values [28%, 83.4%]).

4.3. Children with “below Norm” language skills from study 1

The number of children tested on IEPs was unfortunately quite small, so these results are supplemented with an examination of a subgroup of children from Study 1, whose scores on Vocabulary and Grammar “Product” on QUILS (so as not to include the Fast Mapping module) could be taken to be in the atypical range. Language Ability was included as a fixed factor in our models in Study 1, and at the younger age groups, children’s ability to extend their earlier successful guess increased monotonically as a function of Language Ability. However, for this set of analyses, we partitioned the children from Study 1 into two groups based on their Language Ability. Children who had standard scores in what would be considered the atypical range, taken conservatively to be the bottom 5%, or below a standard score of 75 on either vocabulary or grammar, was grouped as showing “Below-Norm” language skills. These two areas of QUILS do not include the fast-mapping subtest, so the partition is independent. This grouping resulted in a subpopulation of 77 children (40 female). The remainder of the children, namely those scoring above 75, were taken to be “Not-Below-Norm” in language skills. The average vocabulary score of the two groups differed by roughly 20 points, with children in the typical range scoring on average 99.15% (SD = 12.7) on Vocabulary Product, whereas children in the atypical range scored 77.3% (SD = 11.19). We asked whether the two groups differed significantly also in the process by which they learn new words.

As shown in Fig. 14, there is indeed a difference between the two populations. Children who perform in the atypical range converge on the target meaning on the initial trial at much lower rates than those in the typical range, a result that was found to be statistically significant ($b = 0.746$, $SE = 0.103$, $z = 7.279$, $p < 0.001$). That is, they are less able to utilize the available information, e.g. mutual exclusivity and linguistic cues, to home in on the right referent.

However, the two populations differ only in their abilities in extending their choice-of-target on Trial 1 to the second trial (Fig. 15). After a successful Trial 1, children with lower language abilities chose the target object at lower rates on Trial 2 than those with higher skills. However, there does not seem to be a difference between the two

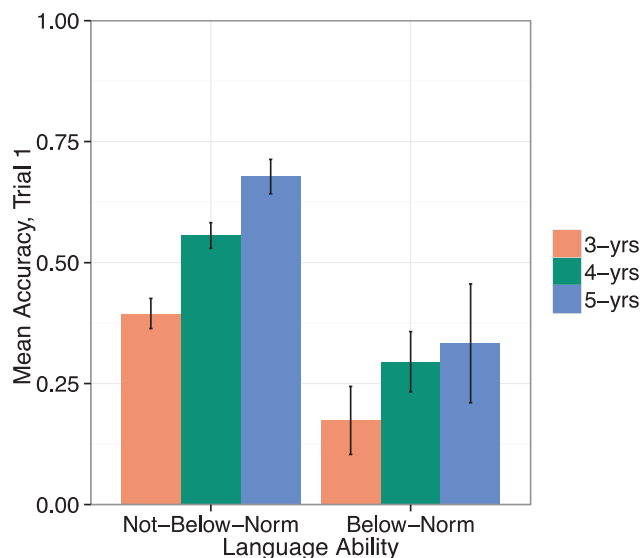


Fig. 14. Accuracy on initial trial by language ability.

populations in rate of success after a failure on Trial 1, with the performance of both groups hovering around chance. Statistical analyses confirm these general trends: A mixed-effects logistic regression with Trial 2 accuracy as our dependent measure and Trial 1 accuracy, Age and Group (below-norm, not-below norm) as fixed factors revealed significant interactions of: (i) Trial 1 and Age Group ($b = 0.22$, $SE = 0.069$, $z = -3.205$, $p = 0.001$), with performance at age 5 showing significant improvement if Trial 1 was successful, and more relevantly, (ii) Trial 1 and Language Ability ($b = 0.23$, $SE = 0.088$, $z = -2.2634$, $p = 0.001$), with children who fall within the typical range for Language Ability performing better, but again, only if their Trial 1 was a success.

To ensure that we were not under-estimating ability, and that the chance-level performance after Trial 1 failure was not driven by low performance from children with low overall language ability, we also asked if performance after Trial 1 failure of the subgroup of children in the typical range was above chance. Results of both the Wilcoxon test ($V = 440373$, $p = 0.9996$) and the intercept-only mixed-effects model ($b = -1.05253$, corresponding to 25.9%, $SE = 0.74$, 95% CI = [−1.198 − 0.906], corresponding to [23.2% 28.8%]) confirm that Trial 1 failers who fall within the typical range are not better than chance on Trial 2.

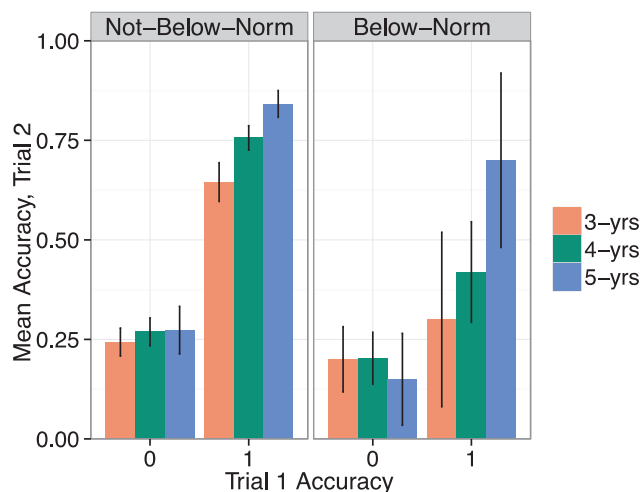


Fig. 15. Trial 2 success by language ability.

4.4. Discussion

Children with language delays were found to have significant difficulties identifying the target referent in the initial trial, but showed the same overall patterns as typical children in our test in their behavior after success vs. failure. Children who made a correct guess initially continued to make the right choice again in Trial 2, though overall, they seemed less able to do so than children in the typical range, especially at the younger ages. However, the two populations are no different after an incorrect guess: neither group retains information about alternative candidates, choosing at random on Trial 2.

5. General discussion and conclusions

5.1. Summary and key observations

In three studies examining word learning in three different populations of child learners, we showed that participants retain information only about their initial hypothesis. Information pertaining to alternative candidates is not retained and used in later instances, even when the second encounter immediately followed the first. When participants guessed correctly on the first trial, they were successful again on Trial 2, though the rate of success was influenced by factors such as age and overall language ability. However, when they guessed incorrectly, they were at chance when selecting among potential referents on Trial 2, suggesting that they did not remember having seen the target referent in the previous encounter. Furthermore, we found no effect of age or language skills on participants' performance on Trial 2 after a failure on Trial 1: irrespective of age and ability, participants do no better than chance.

5.2. Implications for models of word learning

Our results strongly support Propose but Verify models of word learning: children seem to employ a single-hypothesis testing procedure, remembering nothing but their first hypothesis on subsequent learning trials. This was so even though the second learning trial immediately followed the initial trial, and the child's search was guided also by linguistic cues: if the child got the first guess wrong, they do not recover on Trial 2 by recognizing a missed alternative.

While it is clear that these results do not conform to predictions of globalist cross-situational learning models, it is worth considering whether they can be accounted for on hybrid approaches. On [Yurovsky and Frank's \(2015\)](#) model, for instance, a majority portion of attention is allocated towards a single hypothesis and the rest is evenly distributed among the remaining alternatives. The expectation is that as the number of alternative candidates goes up, the likelihood of remembering any one of them decreases, such that after a certain threshold, participants will effectively be at chance after an initial wrong guess. With 4 referent choices, adult participants were able to select a previously seen candidate even after a wrong guess at above-chance rates, suggesting that there were some cross-trial savings. This is not, however, the pattern that we see with children. Of course, this adult-child difference could be an artifact of children's more limited cognitive resources: perhaps children's attention and memory are taxed by 4-referent situations, and preclude strong enough representations of the alternatives as a result. But even if so, we might expect that rates of success after an initial failure should increase with age, as memory and attention also improve as children grow older. We do see a clear improvement with age in the ability to retain initial hypothesis, suggesting that the test is sensitive to potentially age-related differences. However, there was no evidence that children got better at retaining information about alternatives between ages 3 and 6.

5.3. Different types of learners

One of our goals in this paper was to examine and compare the mechanisms of word-learning across different types of learners. Our results reveal both interesting differences as well as striking parallels across populations. Overall, typical monolingual learners, typical dual language learners and atypical learners all show the same general trends in what they carry over from one learning instance to the next. All children seem to succeed after an initial success, but resort to guessing at random after an initial failure, suggesting that they remember only their first hypothesis. This uniformity suggests that children's core cross-situational learning strategy does not shift on the basis of ability or environment. Where the populations do differ is in (i) the ability to utilize the perceptual and linguistic cues available to identify the correct referent initially and (ii) the ability to retain their initial correct guess to follow-up instances.

Consider again the case of dual language learners. They are in a different position with respect to cross-situational learning than monolingual learners and may have to develop a different set of word learning skills, constraints, strategies, and expectations, etc. For instance, the mutual exclusivity assumption cannot help a child who must eventually learn two labels for a given referent, e.g. *dog* and *perro* for dog. Given that the child must exploit mutual exclusivity in order to choose the right referent initially in our task, our finding that dual language learners perform at lower rates than monolinguals on Trial 1 is perhaps unsurprising. Once they have settled on a hypothesis, however, their behavior is no different from monolinguals: they retain no information about alternative candidates. Note that on Propose-but-Verify models, hypotheses about simple concrete words are unidirectional. That is, when a word is heard, a certain referent ought to be available to the senses, but the reverse implication is not true, because a referent that is available to the senses may not be mentioned. There must be many more occasions where, say, a dog is visible but goes unremarked. In the case of dual language learners, the presence of a dog might occasion either *perro*, or *dog*, or go unremarked. Hearing *perro* in the presence of dog does not reduce the likelihood that *dog* is appropriate too, any more than silence or "Rover" blocks it for a monolingual learner. The strong evidence that vocabulary learning is often distributed in the two languages of a dual language learner comes largely from production data and spontaneous speech, not comprehension data. As a result, it is not possible to tell if it arises because the frequency of using certain words in each language is different in context (as is claimed), or whether a kind of mutual exclusivity is operating in bilinguals ([Kalashnikova et al., 2014](#); [Marinova-Todd, 2012](#)). In the former case, it would mean that *perro*, for example, is more often talked about in Spanish, say at Grandma's house, than in English at daycare. In the latter case, it would mean that a child hearing *dog* would block the referent "dog" because it already has the name *perro*. Note that this is not the finding in the work of [Kandhadai et al. \(2017\)](#) with bilingual infants. In that study the infants were more likely than the monolinguals to assign a new word to a category that already had a name. Although accumulating sufficient positive evidence for each set of vocabulary words may certainly take more time for a dual language learner, the underlying mechanisms of learning might not differ.

Finally, we found that language-delayed learners also exhibit the same patterns of behavior as typical learners in what information is retained across trials. Atypical children have greater difficulties in making the right choice on the initial trial and even when they do so, reliably extending their initial right hypothesis to new environments. However, they don't differ from typical children after a failed initial guess: both groups are at chance. These results give us preliminary indications about sources of language delay in vocabulary development. Language-delayed children may have difficulties using word-learning constraints (e.g. mutual exclusivity), contextual and linguistic cues, etc. to extract the right meaning of the word in question. However, the learning procedure itself is a constant: adopt a single hypothesis, and stick with it until it is disconfirmed.

Supplementary material

Data used in this paper can be found in the version control repository for this paper at: <https://github.com/a-aravind/WordLearning>.

Funding

This work was supported by IES grant R305A110284.

References

- Byers-Heinlein, K., & Werker, J. (2009). Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word-learning heuristic. *Developmental Science*, *12*, 815–823.
- Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Papers and Reports on Child Language Development*, *15*, 17–29.
- Dautriche, I., & Chemla, E. (2014). Cross-situational word learning in the right situations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 892–903.
- Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Journal of Developmental Psychology*, *28*(1), 99–108.
- Golinkoff, R., De Villiers, J. G., Hirsh-Pasek, K., Iglesias, A., & Wilson, M. S. (2017). *User's manual for the Quick Interactive Language Screener ((QUILS(TM)) a measure of vocabulary, syntax, and language acquisition skills in young children*. Baltimore: Md. Paul H. Brookes Publishing Co.
- Grosjean, F. (2008). *Studying bilinguals*. Oxford: Oxford University Press.
- Grosjean, F. (2010). *Bilingual: Life and reality*. Cambridge, MA: Harvard University Press.
- Hollich, G., Golinkoff, R. M., & Hirsh-Pasek, K. (2007). Young children associate novel words with complex objects rather than salient parts. *Developmental Psychology*, *43*, 1051–1061.
- Iglesias, A., de Villiers, J.G., Golinkoff, R., Hirsh-Pasek, K., & Wilson, M.S. (in press). The QUILS-ES: a language screener for young dual language learners of English and Spanish. Baltimore: Brookes Publishing Company.
- Kalashnikova, M., Mattock, K., & Monaghan, P. (2014). Disambiguation of novel labels and referential facts: A developmental perspective. *First Language*, *34*, 125–135.
- Kandhadai, P., Hall, G., & Werker, J. (2017). Second label learning in monolingual and bilingual infants. *Developmental Science*, *20*. <http://dx.doi.org/10.1111/desc.12429>.
- Kaushanskaya, M., & Marian, V. (2009). The bilingual advantage in novel word learning. *Psychonomic Bulletin & Review*, *16*, 705–710.
- Markman, E. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*(1), 57–77.
- Markman, E. (1989). *Categorization and naming in children: Problems of induction*. Cambridge: MIT Press.
- Marian, V., & Shook, A. (2012). The cognitive benefits of being bilingual. *Cerebrum: The Dana forum on brain science* (Vol. 2012). Dana Foundation.
- Marinova-Todd, S. H. (2012). “Corplum is a core from a plum”: The advantage of bilingual children in the analysis of word meaning from verbal context. *Bilingualism: Language and Cognition*, *15*(1), 117–127.
- 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*, 831–877.
- Medina, T. N., Snedeker, J., Trueswell, J. C., & Gleitman, L. R. (2011). How words can and cannot be learned by observation. *Proceedings of the National Academy of Sciences*, *108*, 9014–9019.
- Poepsel, T. J., & Weiss, D. J. (2016). The influence of bilingualism on statistical word learning. *Cognition*, *152*, 9–19.
- Quine, W. V. O. (1960). *Word and object*. Cambridge, MA: MIT Press.
- Roembke, T. C., & McMurray, B. (2016). Observational word learning: Beyond propose-but-verify and associative bean counting. *Journal of Memory and Language*, *87*, 105–127.
- Siskind, J. (1996). A computational study of cross-situational techniques for learning word-to-meaning mappings. *Cognition*, *61*(1), 39–91.
- Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition*, *106*(3), 1558–1568.
- Snedeker, J., & Gleitman, L. (2004). Why it is hard to label our concepts. In D. G. Hall, & S. R. Waxman (Eds.). *Weaving a lexicon* (pp. 257–294). Cambridge, MA: MIT Press.
- Stevens, J., Gleitman, L., Trueswell, J., & Yang, C. (2017). The pursuit of word meanings. *Cognitive Science*, *41*(4), 638–676.
- Trueswell, J., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets cross-situational word learning. *Cognitive Psychology*, *66*(1), 126–156.
- Vlach, H., & Debrock, C. (2017). Remember dax? Relations between children's cross-situational word learning, memory, and language abilities. *Journal of Memory and Language*, *93*, 217–230.
- Vouloumanos, A. (2008). Fine-grained sensitivity to statistical information in adult word learning. *Cognition*, *107*, 729–742.
- Woodard, K., Gleitman, L., & Trueswell, J. (2016). Two-and three-year-olds track a single meaning during word learning: Evidence for propose-but-verify. *Language Learning and Development*, *12*, 252–261.
- Yu, C., & Smith, L. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*(5), 414–420.
- Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*(2), 244–262.
- Yurovsky, D., & Frank, M. C. (2017). Beyond naïve cue combination: Salience and social cues in early word learning. *Developmental Science*, *20*(2).
- Yurovsky, D., Fricker, D. C., Yu, C., & Smith, L. B. (2014). The role of partial knowledge in statistical word learning. *Psychonomic Bulletin & Review*, *21*, 1–22.
- Yurovsky, D., & Frank, M. C. (2015). An integrative account of constraints on cross-situational learning. *Cognition*, *145*, 53–62.