



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Prediction error boosts retention of novel words in adults but not in children

Citation for published version:

Gambi, C, Pickering, MJ & Rabagliati, H 2021, 'Prediction error boosts retention of novel words in adults but not in children', *Cognition*, vol. 211, 104650. <https://doi.org/10.1016/j.cognition.2021.104650>

Digital Object Identifier (DOI):

[10.1016/j.cognition.2021.104650](https://doi.org/10.1016/j.cognition.2021.104650)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Cognition

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



PREDICTION ERROR AND WORD LEARNING

1 Prediction error boosts retention of novel words in adults but not in children.

2 Chiara Gambi

3 University of Edinburgh and Cardiff University

4 Martin J. Pickering

5 Hugh Rabagliati

6 University of Edinburgh

7

8 Address for correspondence:

9 Chiara Gambi

10 School of Psychology

11 70, Park Place

12 Cardiff University

13 CF10 3AT Cardiff, U.K.

14 GambiC@cardiff.ac.uk

15 Phone: +44(0)29 206 88950

16

17 --- Manuscript accepted for publication in *Cognition* ---

18

19

20

21

22

23

24

25

26

Abstract

How do we update our linguistic knowledge? In seven experiments, we asked whether error-driven learning can explain under what circumstances adults and children are more likely to store and retain a new word meaning. Participants were exposed to novel object labels in the context of more or less constraining sentences or visual contexts. Both two-to-four-year-olds ($M_{\text{age}} = 38$ months) and adults were strongly affected by expectations based on sentence constraint when choosing the referent of a new label. In addition, adults formed stronger memory traces for novel words that violated a stronger prior expectation. However, preschoolers' memory was unaffected by the strength of their prior expectations. We conclude that the encoding of new word-object associations in memory is affected by prediction error in adults, but not in preschoolers.

Keywords: prediction error; mutual exclusivity; disconfirmed predictions; memory retention; word learning.

PREDICTION ERROR AND WORD LEARNING

60 Prediction error boosts retention of novel words in adults but not in children.

61 Children learn new words at a staggering rate (Fenson et al., 1994), demonstrating a
62 remarkable ability not only to determine what a new word means, but also to retain huge numbers
63 of form-meaning pairs in memory (Vlach, 2019). This learning extends into adulthood and indeed is
64 lifelong, as new terms and vocabulary enter our language, and as we move between different
65 linguistic communities (Ameel, Malt, & Storms, 2008; Borovsky, Kutas, & Elman, 2010; Hulme,
66 Barsky, & Rodd, 2019). In this work, we investigate what factors affect our ability to retain word
67 meanings, and whether these are the same in children and adults. In particular, we test how the
68 retention of novel form-meaning pairs is affected by prediction errors, following theoretical claims
69 that the computation of prediction errors drives memory encoding (Henson & Gagnepain, 2010).

70 There is now ample evidence that our interactions with the world are guided by prediction,
71 from the way we control our movements (Wolpert & Flanagan, 2001) to how we make sense of our
72 perceptions (Clark, 2013; Friston, 2005; Grush, 2004). Across these different domains, we are able
73 to generate expectations about the future state of the world and, critically, we compare these
74 expectations to information about the actual state of the world when it reaches our senses. This
75 process of comparison between expected and observed states generates prediction error signals,
76 which are thought not only to drive immediate behavioral responses, but also to affect long-term
77 encoding of information in memory (Henson & Gagnepain, 2010), and thus our learning (e.g., Den
78 Ouden, Friston, Daw, McIntosh, & Stephan, 2008; Niv & Schoenbaum, 2008; Rescorla & Wagner,
79 1972).

80 Importantly, prediction error is the result of a comparison between expected and observed
81 states, and thus its magnitude depends on the strength, or precision, of both the information we
82 receive from the outside world and of our prior expectations (Friston, 2005, 2010). Under the
83 Predictive Interactive Multiple Memory Systems (PIMMS) framework proposed by Henson and
84 Gagnepain (2010), larger prediction errors (i.e., greater mismatches between expected and observed
85 states) lead to the formation of stronger memory traces. Combined with the idea that stronger (i.e.,

86 more precise) expectations generate larger prediction errors (when disconfirmed), this leads to a key
 87 hypothesis: Stronger expectations that are disconfirmed should benefit memory more than weaker
 88 expectations that are disconfirmed. While this may seem surprising and even counterintuitive (after
 89 all, incorrect expectations are akin to mistakes and making mistakes should impair memory for the
 90 correct answer), it falls out of the way prediction error is defined in these accounts – that is, as the
 91 discrepancy between expectations and input.

92 For example, Greve and colleagues showed that adults were more likely to remember the
 93 association between a scene and a new face (observed only once) if the scene had previously been
 94 repeatedly paired with another face (i.e., the same face multiple times), compared to several faces
 95 all different from the new face (Greve, Cooper, Kaula, Anderson, & Henson, 2017). Crucially,
 96 although in both instances the new face violated a previously established association (i.e., it
 97 disconfirmed an expectation), the previously-established association supported a stronger
 98 expectation when the scene was paired repeatedly with the same face. Thus, this finding confirms a
 99 key hypothesis derived from accounts of memory based on the computation of prediction error.

100 But does prediction error also affect our memory for word meanings? Surprisingly, despite a
 101 lot of recent interest in adults' and children's ability to predict upcoming language (Huettig, 2015;
 102 Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018; Pickering & Garrod, 2013; Rabagliati,
 103 Gambi, & Pickering, 2016), the answer to this question is still unclear. While much evidence
 104 demonstrates that adults and young children are capable of generating expectations at multiple
 105 linguistic levels (Pickering & Gambi, 2018; Rabagliati et al., 2016), including meaning (Altmann &
 106 Kamide, 1999; Borovsky, Elman, & Fernald, 2012; Lindsay, Gambi, & Rabagliati, 2019; Mani,
 107 Daum, & Huettig, 2016; Mani & Huettig, 2012), structure (Gambi, Pickering, & Rabagliati, 2016;
 108 Havron, de Carvalho, Fiévet, & Christophe, 2019; Lukyanenko & Fisher, 2016; Wicha, Moreno, &
 109 Kutas, 2004), and perhaps form (Dikker, Rabagliati, Farmer, & Pylkkänen, 2010; Ylinen et al.,
 110 2014; but see Gambi, Gorrie, Pickering, & Rabagliati, 2018), comparatively little work has
 111 examined the consequences of disconfirmed expectations on memory for novel word meanings.

112

113 **Prediction Error and Language Learning**

114 The idea that prediction errors influence language learning is not new, and indeed a number of
115 historically important models have made the claim that prediction errors play a critical role in
116 children's language development. In these models, prediction error typically acts as a guide for
117 learning; it offers a signal for when the learner should (or should not) revise their linguistic
118 knowledge. For instance, Elman (1990; see also St. John & McClelland, 1990) introduced the idea
119 that prediction error-driven learning could help a simple recurrent connectionist network acquire
120 approximate linguistic representations: The network was trained to predict the next word in a large
121 corpus of text and, when an encountered word mismatched its prediction, the model's internal
122 representations were revised through backpropagation of error (Rumelhart, Hinton, & Williams,
123 1986). These ideas have also been highly influential for newer models of grammatical development
124 (e.g., Chang, Dell, & Bock, 2006; Dell & Chang, 2014) and word learning (Plaut and Kello (1999)).
125 In addition, related ideas about error-driven learning can be seen in models that use theories of
126 reinforcement learning to explain language development. For instance, Ramscar, Yarlett, Dye,
127 Denny, and Thorpe (2010) argued that a model based on the Rescorla-Wagner learning rule
128 (Rescorla & Wagner, 1972) can capture how children acquire word meanings under conditions of
129 referential uncertainty, because the computation of prediction errors allows the child to discriminate
130 between the situations in which a word can or cannot be used (see also Ramscar, Dye, & McCauley,
131 2013b). Thus, across all of these models, prediction errors guide children in forming linguistic
132 representations that can accurately predict the linguistic input that they are likely to encounter.

133 In the PIMMS framework (described above), prediction errors also guide learning, but they
134 do so by indexing how robustly the learner should encode a piece of encountered information into
135 memory. Specifically, unexpected information (i.e., information that generates a larger prediction
136 error) is encoded more strongly and thus can be retrieved more easily in the future. For the task of
137 learning a word, this framework highlights that prediction errors could influence how learners

138 remember and retain word meanings over longer periods of time. While this is not fundamentally
 139 different from the models reviewed above, in those models the focus is on how learners discover the
 140 meanings of words through experience (for examples see Ramscar et al., 2010; Grimmick,
 141 Gureckis, & Kachergis, 2019; Stevens, Gleitman, Trueswell, & Yang, 2017): Prediction errors
 142 generated by current input guide changes in linguistic representations and ensure that the system
 143 can accurately predict future input. The PIMMS framework instead focuses on prediction error's
 144 influence on retention of novel information in memory, which is the topic we address here, in both
 145 adults and children.

146 At least since Carey and Bartlett (1978), it has been recognised that young children can
 147 accurately retain word meanings in long-term memory, though exactly how much they are able to
 148 retain and under what conditions has been debated (e.g., Horst & Samuelson, 2008; Spiegel &
 149 Halberda, 2011; Vlach & Sandhofer, 2012; see Samuelson & McMurray, 2017, for review). In any
 150 case, to the extent that children do retain word meanings, this long-term retention appears to rely on
 151 domain-general memory mechanisms (Markson & Bloom, 1997, Vlach, 2019; Vlach & DeBrock,
 152 2017). For instance, children's ability to retain word meanings up to one month is roughly matched
 153 to their ability to retain non-linguistic factual information over the same length of time (Markson &
 154 Bloom, 1997; Vlach & Sandhofer, 2012), and their memory for word meanings is affected by
 155 factors that are known to influence memory for non-linguistic information, such as repetition and
 156 spacing (e.g., Sandhofer & Vlach, 2011) and sleep (e.g., Henderson, Weighall, Brown, & Gaskell,
 157 2012). While adults' retention rates for novel word meanings can be higher than children's, they are
 158 similarly matched to their retention rates for novel non-linguistic information (Markson & Bloom,
 159 1997; Sandhofer & Vlach, 2012).

160 Given these considerations, and the findings that prediction errors influence memory for
 161 non-linguistic information in adults, we might expect prediction error should also affect the
 162 retention of word meanings over time, such that retention accuracy is greater for words that are

163 learned in unexpected contexts. However, the prior evidence for this is actually somewhat unclear,
164 as we review below.

165 **Adults.** We are not aware of any study that has tested how prediction error affects retention
166 of newly-learnt word meanings in adults. However, a small number of studies do provide indirect
167 evidence for a role of prediction error in adult word learning. Fitneva and Christiansen showed that
168 adults perform better in a learning task when they encounter a greater proportion of word-referent
169 mappings that are unexpected (Fitneva & Christiansen, 2011, 2017). Using a cross-situational
170 learning paradigm, where novel words are repeatedly presented under situations of referential
171 ambiguity (i.e., with multiple potential referents for each word; Yu & Smith, 2007), they exposed
172 learners to word-referent mappings that were unexpected because they differed from those trained
173 during an initial familiarization phase. Other mappings were instead expected, as they did not differ
174 from those established during the familiarization phase. Strikingly, when the proportion of
175 unexpected to expected mappings was higher, adult learners actually learned more compared to
176 when the proportion of unexpected mappings was lower.

177 But while this finding may suggest that prediction error plays a role in adult word learning,
178 it is unclear whether this interpretation is correct. According to a prediction error account,
179 participants generated expectations about the words they were going to hear based on the mappings
180 established during the familiarization phase and, the more often these expectations were then
181 disconfirmed (because many mappings had changed), the more the resulting error signals benefitted
182 learning of new word-referent pairings. But if this advantage stems from prediction errors, then it
183 should be specific to the unexpected mappings – because it is only for these items that the learner
184 should generate incorrect expectations. In contrast, Fitneva and Christiansen (2011, 2017) found
185 that both unexpected and expected mappings were learned better when the proportion of unexpected
186 mappings was higher, suggesting a very different explanation: The larger number of errors may
187 have prompted participants to allocate more attentional resources to the task, and thus process all
188 words and referents more deeply. However, since a recent study instead found that learning was

189 enhanced specifically for unexpected mappings (Grimmick et al., 2019), it remains possible that
190 prediction errors do play a role in adult word learning.

191 **Children.** More child studies are relevant to our question, but the picture that emerges from
192 them is also mixed. Two strands of work suggest that, far from driving learning, generating
193 incorrect expectations may hinder children’s processing of new information, with negative
194 consequences for their ability to learn this information. First, when Fitneva and Christansen (2017)
195 asked whether 4-year-olds’ word learning would benefit from encountering high proportions of
196 unexpected mappings (as in adults), they instead found that 4-year-olds learn better when the
197 proportion of unexpected mappings is lower. But since it is unclear whether the findings in adults
198 demonstrate a role for prediction error, it is also unclear whether the findings in children provide
199 evidence against it, for the same reason: The expectancy-driven effects were not item-specific.

200 Second, Benitez and colleagues showed that infants (Benitez & Smith, 2012) and 2-year-
201 olds (Benitez & Saffran, 2018) learn novel word-referent associations better when the associations
202 are demonstrated in predictable contexts, compared to unpredictable contexts. While this seems to
203 go against the idea that prediction error boosts learning, in this task predictability was manipulated
204 by having the referents appear in predictable or unpredictable spatial locations, but the association
205 between the words and their referents were not themselves more or less predictable. Therefore, the
206 prediction error signal may have enhanced memory, but for the location of the stimuli (which was
207 not tested), rather than for the word-referent association.

208 In contrast, other evidence suggests that disconfirmed expectations boost children’s
209 memory. First, Stahl and Feigenson (2017) showed that 3-to-6-year-olds are more likely to
210 remember a novel action word if the action it refers to is unexpected due to violations of physical
211 “core knowledge” (e.g., a bag “magically” changing the color of objects that are put inside it).
212 However, in the control condition where no expectation was violated (i.e., the object behaved
213 “normally”), children did not learn the novel word at all (they performed at chance), likely because
214 in this condition there was no salient action, and the very use of a novel word was thus

215 pragmatically infelicitous. This finding suggests that perhaps the unexpected action did not boost
216 memory because it violated an expectation, but rather because it created the pragmatic conditions
217 for use of a novel word. Moreover, actions that violate core knowledge are not just unexpected, but
218 outright impossible, so this conclusion may not generalize to word learning in the wild.

219 Second, potential evidence that children learn from disconfirmed expectations comes from
220 Reuter, Borovsky, and Lew-Williams's (2019) eye-tracking study. Three-to-five-year-olds heard
221 novel words while observing two potential referents, one of which was a familiar object whose
222 name was likely known to the child, while the other was a novel (and thus nameless) object. Infants
223 as young as 16 months (Halberda, 2003; Horst & Samuelson, 2008) reliably map a novel word unto
224 the novel object at first exposure under these conditions, following the so-called mutual exclusivity
225 constraint. Crucially, Reuter and colleagues embedded the novel words within sentences, and
226 manipulated the degree of semantic constraint of such sentences, so that they would provide either a
227 strong expectation for the name of the familiar object (high constraint) or no strong expectation
228 (low constraint). For example, a child looking at pictures of a spoon and a novel object should
229 generate a strong expectation of *spoon* following *Yummy! Let's eat soup. I'll stir it with a...*,
230 whereas following *Neat! Look over there. Take a look at the...*, no strong expectation for either
231 object should be generated.

232 The child then heard a novel word (e.g., ...*cheem*) at the end of both high and low constraint
233 sentences. As a result of the expectation-strength manipulation, the novel word disconfirmed a
234 stronger prior expectation in the high than in the low constraint condition, thereby generating a
235 larger prediction error signal in the former than the latter condition. Reuter and colleagues
236 hypothesized that novel words associated with larger prediction errors should be better learnt, and
237 used a preferential looking task to test this: They presented children with each novel word and two
238 novel referents (the target, and a distractor that was the correct referent for a different novel word),
239 and measured whether the child looked more at the correct referent than the distractor. Surprisingly,
240 children's performance was at chance with words encountered after high-constraint sentence

PREDICTION ERROR AND WORD LEARNING

241 contexts (and instead above chance in the low constraint condition). This finding is difficult to
242 reconcile with prediction error being the driver of children’s memory: If it were, children should
243 have been more likely to gaze at the correct referent in the high than the low constraint condition,
244 because novel words disconfirmed a stronger prior expectation in the former than the latter
245 condition.

246 Nevertheless, Reuter et al. (2019) suggested that their findings support error-driven accounts
247 of novel word learning because they also found a positive correlation between each child’s ability to
248 revise following a disconfirmed prediction and their performance at test. Specifically, they
249 computed a “predict-and-revise” looking measure, which was larger the more the child looked at
250 the familiar object before hearing the novel word (i.e., the stronger their prior expectation) *and* the
251 more they looked towards the novel object upon hearing the novel word (i.e., the faster they revised
252 their prior expectation). They argued that a positive correlation between children’s “predict-and-
253 revise” looking pattern during learning and the extent to which they preferentially looked at the
254 target referent for the novel word at test was evidence that the revision of incorrect expectations
255 was driving learning. But while this correlation was specific to novel words that were embedded in
256 high-constraint sentences (i.e., no correlation was present for items that were presented at the end of
257 low-constraint sentences), this interpretation is at odds with the lack of an overall memory
258 advantage for words presented in the high-constraint condition.

259 Thus, we suggest an alternative interpretation is more likely: High-constraint sentences may
260 have hampered memory by shifting attention away from the novel object, and only those children
261 who were able to recover from this attentional shift would have learned the correct referent for the
262 novel word. In contrast, low-constraint sentences did not reduce children’s preference for looking at
263 the novel object, thus supporting memory regardless of individual differences in sentence
264 processing ability. Importantly, under this interpretation, the relation between individual children’s
265 learning and their sentence processing ability could be entirely explained by a common underlying

266 factor, such as processing speed, rather than being explained by a specific ability to predict-and-
267 revise.

268 In sum, we do not know whether a prediction-error mechanism underlies the formation and
269 consolidation of novel word-object associations: Theories of memory based on prediction error
270 predict that adults and children should form stronger memory representations following the
271 disconfirmation of stronger expectations, but the empirical evidence is inconclusive. There is
272 evidence that children have weaker memories for novel words when they are encountered in
273 unexpected contexts (Benitez & Smith, 2012; Benitez & Saffran, 2018), but also some suggestive
274 evidence that strong but incorrect expectations may in fact be beneficial (Stahl & Feigenson, 2017;
275 Reuter et al., 2019). In adults, no study has contrasted memory following a stronger than a weaker
276 disconfirmed expectation.

277 Thus, our first question is whether adults acquire stronger memories for new word-object
278 associations, if they are observed in the context of a violation of a stronger linguistic expectation.
279 Since word learning is a lifelong process, our second question is whether the underlying
280 mechanisms remain similar across the lifespan or whether they themselves develop, and therefore
281 we also compare adult performance to that of 2-to-4-year-olds on the same learning task. If we find
282 evidence that this effect emerges early in development, this would suggest that the computation of
283 prediction errors plays a role in word learning from the early stages of language acquisition.

284 **The current study**

285 Using a task similar to Reuter et al. (2019), we asked whether expectation strength affects
286 the strength of memory representations for the mapping between a novel word and its referent.
287 Importantly, we did so both for children (2-to-4-year-olds) and young adults (university students),
288 so we could directly observe any developmental changes in the mechanisms used for word learning
289 (unlike Reuter et al., who tested only children). While we disagree with Reuter et al.'s interpretation
290 of their findings, note that we do not take issue with their design, and in fact we adopt a very similar
291 design.

292 Reuter et al.'s (2019) design has two key strengths. First, it allows for a comparison between
293 disconfirmed expectations that differ in strength (of the a priori expectation): This is an ideal
294 comparison for testing the effect of prediction errors on subsequent memory (see Greve et al.,
295 2017). In contrast, most previous studies (Stahl & Feigenson, 2017; Benitez & Smith, 2012; Benitez
296 & Saffran, 2018) compared confirmed to disconfirmed predictions, which is problematic because
297 these conditions do not just differ in the magnitude of the prediction error: When a prediction is
298 disconfirmed, the predicted but not encountered word may linger in memory (Rommers &
299 Federmeier, 2018), and potentially counteract the benefits of a larger prediction error on memory, if
300 it interferes with encoding of observed word. Second, expectations strength is manipulated using
301 sentence contexts (rather than artificially, by changing word-referent mappings mid-way through
302 the experiment, as in Fitneva & Christiansen, 2011, 2017), thus providing a more ecologically valid
303 test. We retained both of these aspects of Reuter et al.'s (2019) study, but our procedure did differ
304 from theirs in some important respects, which we highlight below.

305 Like Reuter et al. (2019), we manipulated contextual constraint in order to vary expectation
306 strength: Adult and child learners encountered novel words (e.g., *cheem*) embedded within
307 sentences that were either more constraining (*Now, Peppa will eat the cheem*) or less constraining
308 (*Now, Peppa will get the cheem*) with respect to the visual context. The visual context always
309 consisted of two objects: a familiar object that fit the more constraining verb (e.g., an apple for *eat*)
310 and an unfamiliar object (e.g., the jelly-like object in Figure 1). Since the younger children we
311 tested were 2-year-olds (vs. 3-year-olds in Reuter et al., 2019), we took our constraining verbs from
312 Mani and Huettig (2012), who showed evidence for prediction in 24-month-olds.

313 Based on the vast literature on linguistic prediction in adults and children, we expected
314 listeners to generate a strong expectation that the familiar object would be mentioned following an
315 High Constraint context – because Peppa is much more likely to eat the apple than the jelly-like
316 object in Figure 1 (as we confirmed in a post-test; see Methods below). In contrast, following the
317 Low Constraint context, listeners could generate only a weaker expectation (or no expectation at

318 all), because “Now, Peppa will get the...” is not as strongly predictive of “apple” as “Now, Peppa
319 will eat the...” is.

320 In order to disconfirm listeners’ expectations, we relied on the presence of the unfamiliar
321 object in the display (similarly to Reuter et al., 2019). Since even very young children prefer to map
322 novel words onto unfamiliar (and thus nameless) objects when the alternative is a familiar object
323 with a known name (i.e., they follow the mutual exclusivity constraint; Halberda, 2003), both
324 children and adults should be biased to revise their expectations, and select the unfamiliar object as
325 the referent of the novel word. This bias may not operate on 100% of trials, so sometimes
326 participants may select the familiar object as a referent for the novel word. In such cases, it may be
327 that participants noticed the novel word but chose to interpret it as a novel name for the familiar
328 object (e.g., the name of a novel variety of apple), or it may be that they failed to notice the novel
329 word (e.g., because they followed their expectations).

330 Because it is hard to discriminate between these two options, we conducted analyses that
331 exclude such cases, and are restricted to instances in which participants selected the unfamiliar
332 object explicitly, as in these cases we can be certain that they mapped the novel word onto the
333 unfamiliar object. Crucially, while the occurrence of the novel word should disconfirm participants’
334 expectations following both more and less constraining contexts, the magnitude of the resulting
335 prediction error should be larger following more constraining contexts, where the strength of the
336 prior expectation was higher. Compare Figures 1a and 1b for a graphical illustration of the
337 processes at play during high and low constraint learning trials.

338 To test how memory depends upon processing during learning, we asked participants to
339 select a referent for each novel word at test to probe retention of the novel mappings. Note that this
340 task differs from the preferential looking measure used by Reuter et al. (2019), and it is a more
341 explicit measure of memory. We chose this explicit measure because it is the one used in much
342 research on the mutual exclusivity constraint. Studies that have tested 2-year-olds on similar tasks
343 have shown that, even though children correctly map the novel word *cheem* to the unfamiliar object

PREDICTION ERROR AND WORD LEARNING

344 on the fly, they often fail to retain the mapping established via mutual exclusivity in memory when
345 tested at short (i.e., on the order of 5-10 minutes) retention intervals (e.g., Horst & Samuelson,
346 2008; see Samuelson & McMurray, 2017 for review; but cf. Spiegel & Halberda, 2011). While this
347 means we were expecting the youngest children to perform well below ceiling overall in our
348 explicit memory test, it also provides an additional motivation for our study: If children initially
349 encode novel words only weakly in memory after a first encounter, is it possible to strengthen such
350 memory traces by encouraging them to generate linguistic expectations that will be later
351 disconfirmed?

352 To summarize, we hypothesize that both adults and young children should be more likely to
353 remember novel words that violate stronger, as opposed to weaker expectations. We test this
354 hypothesis in 7 experiments (see Table 1 for an overview).

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369 Table 1. Overview of experiments. Please refer to the text for an explanation of the
 370 differences between experiments.

Experiment	Participants	Aim	Manipulation	Testing modality / task during break	Context repetitions
1	40 adults	Power calculation	Verb constraint	Experimenter present / tapping, conversation with experimenter	1
2	40 adults	Power calculation	Verb constraint	Online / video + comprehension questions	2
3	58 adults	Replication of Exp. 1-2	Verb constraint	Online / video + comprehension questions	1
4	58 adults	Control experiment	Object distractor	Online / video + comprehension questions	1
5	65 adults	Replication of Exp. 4	Object distractor	Online / video + comprehension questions	1
6	80 children	Child version of Exp. 1-3	Verb constraint	Experimenter present / tapping	1
7	86 children	Child version of Exp. 4-5	Object distractor	Experimenter present / tapping	1

371

372 **Experiments 1-3: Verb-constrained prediction errors in adults**

373 While the paradigm was designed with children in mind, we first tested it on adult participants to
 374 assess the robustness of the effect. We established this in three experiments, which differed
 375 minimally in procedure. These differences are described below (and summarized in Table 1), but
 376 since findings were consistent across experiments, here we present combined results. All materials,
 377 data, and analyses scripts, including separate analyses and graphs for each experiment are available

378 at https://osf.io/zvn6u/?view_only=ce8cc8f5432e41019498a98a2687982b, see Additional analyses
379 in the *analysis_scripts* folder, section 1).

380 **Methods.**

381 **Participants.** Experiments 1 and 2 tested 40 adults each; this sample size was a rough estimate, and
382 it was expected to yield around 80% power only with a large effect size ($d=0.8$; Westfall, Kenny, &
383 Judd, 2014). We then used a simulation approach to compute sample size (N) for subsequent
384 studies. We did this through a bootstrapping approach: we repeatedly (1000 times) randomly
385 sampled N adult participants, analyzed retention accuracy as reported below (Data Analysis), and
386 extracted the z statistics associated with the effect of interest (i.e., the effect of sentence constraint).
387 We defined power as the percentage of samples that yielded z equal to or greater than 1.645 - i.e.,
388 the threshold for significance of a one-tailed test, as our prediction is directional: High Constraint
389 contexts should lead to better memory than Low Constraint contexts. When this procedure was
390 applied to data from Experiments 1 and 2, it indicated that 58 participants would achieve 95%
391 power, so we recruited that many participants for a replication (Experiment 3). In total, 138
392 University of Edinburgh students (32 male, age range: 17 to 31; 40 participants did not provide age
393 information) took part across the three experiments, either for course credit or £2; 16 reported to be
394 native speakers of a language other than English, but since this did not affect the results (see
395 Additional analyses in the *analysis_scripts* folder on the OSF, section 4), the analyses below
396 disregard language status. The study received ethical approval from the University of Edinburgh.

397

398

399

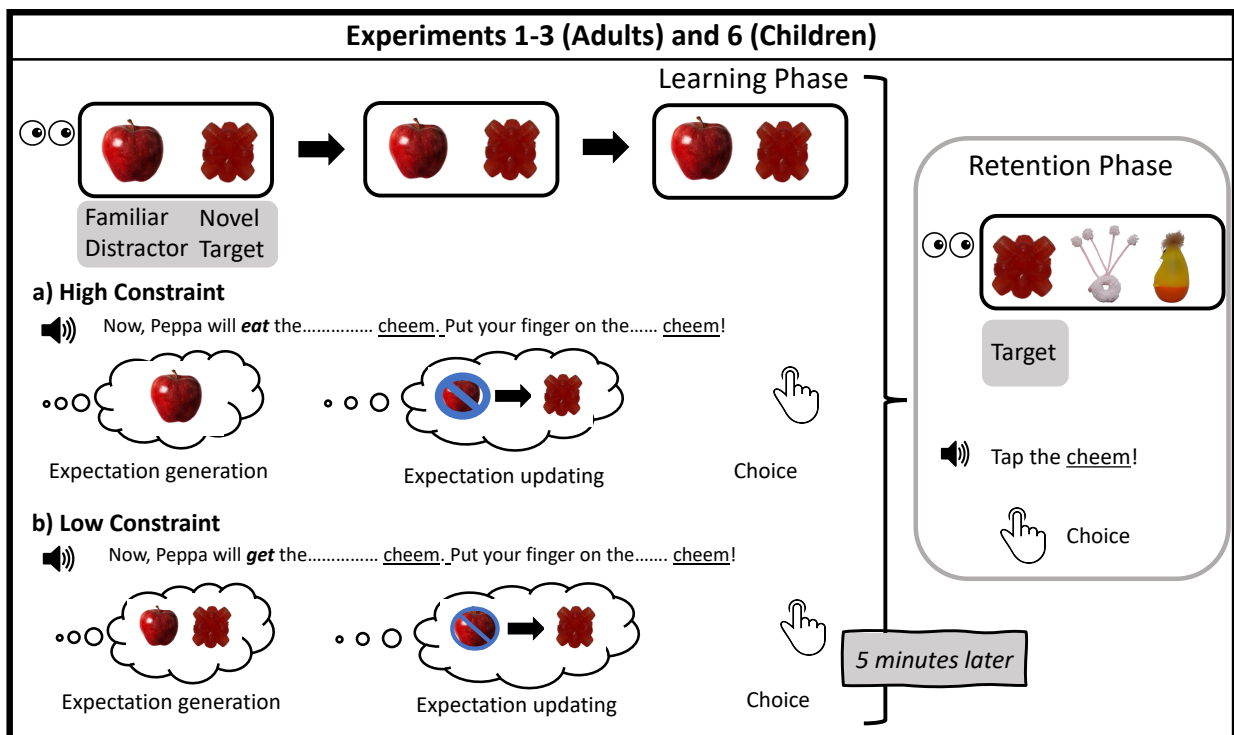
400

401

402

PREDICTION ERROR AND WORD LEARNING

403 Figure 1. Schematic depiction of the experimental design in Experiments 1-3 and 6, including a
404 graphical illustration of the processes at play during different types of learning trials: a) High
405 Constraint trials, b) Low Constraint trials. Note that in (a) we conservatively assume no expectation
406 that the novel object will be named before the participant hears the novel word – this is because the
407 novel objects only had a loose fit with the constraining verb (e.g., the spiky red object in the figure
408 had a jelly-like consistency). We depict only learning trials on which participants choose the novel
409 object as the referent of the novel word.



410
411 **Materials and Procedure.** The experiments consisted of two phases (see Figure 1). In the learning
412 phase (top), participants completed 14 trials: Following two practice trials, 8 experimental trials
413 were randomly interspersed with 4 filler trials. All learning trials had the same structure.
414 Participants saw a picture of the cartoon character Peppa Pig centered on the top half of the screen.
415 On the bottom half of the screen, they saw photographs of a familiar and an unfamiliar object.

416 Participants began a trial by clicking or tapping on the picture of Peppa Pig, which triggered
417 a pre-recorded sentence. To test whether repetition helps participants revise a disconfirmed
418 expectation, in Experiment 2 adults heard two sentences, so the target word was always presented at

PREDICTION ERROR AND WORD LEARNING

419 least twice. However, adults' performance in Experiment 2 did not differ from Experiment 1, where
420 only one sentence was used. Thus, Experiment 3 and all other experiments reported here used only
421 one sentence. Participants could listen to the sentence as many times as they wished by tapping on
422 the top picture again.

423 Filler sentences were always high constraint and mentioned the predictable, familiar object
424 (e.g., *Now, Peppa will rock the baby*) in order to encourage participants to predict familiar words.
425 Crucially, on half the experimental trials participants listened to a High Constraint sentence (e.g.,
426 *Now, Peppa will eat the*, when the familiar object was an apple), but on the other half they listened
427 to a Low Constraint sentence (e.g., *Now, Peppa will get the*). This way we manipulated the degree
428 to which participants expected to hear the name of the familiar object. Constraint was manipulated
429 within participants and items, counterbalanced across two lists.

430 While filler sentences always ended with the name of the familiar object, experimental
431 sentences ended with one of 8 novel pseudowords (*cheem, dite, doop, fode, foo, pabe, roke* and
432 *yok*), mostly drawn from Horst and Samuelson (2008). After the sentence, learners heard an
433 instruction (e.g., *Put your finger on the cheem!*) asking them to select the object corresponding to
434 the final word in the sentence. Unfamiliar objects were selected from Horst and Hout's (2016)
435 NOUN database; familiarity and nameability were kept as low as possible, but such that the novel
436 objects would always match the constraint of the verb in High Constraint sentences (e.g., the object
437 paired with *eat* had to look edible). A post-test with 20 adults (7 males, 22 to 61 years of age)
438 recruited from the online platform CloudFlower confirmed that novel objects were a better fit for
439 the constraining verbs they were paired with ($M = 3.08$ on a 1-to-7 Likert scale), than for another
440 (randomly selected) constraining verb ($M = 2.19$, $t(19) = 4.12$, $p < .001$). The same post-test showed
441 that, unsurprisingly, familiar objects were a better fit for the constraining verbs ($M = 6.10$)
442 compared to the unfamiliar objects ($M = 3.08$) they were paired with ($t(19) = 7.85$, $p < .001$). We
443 return to this issue below as it was part of the motivation for conducting Experiments 4 and 5.

PREDICTION ERROR AND WORD LEARNING

444 Following completion of the learning phase, participants took a short (approximately 5-
445 minute-long) break. What happened during the break depended on whether the experiment was
446 conducted in the lab or online. Participants in Experiment 1 were tested in the lab and, during the
447 break, they first tapped on a series of cartoon characters (this task was designed for children, and is
448 described in more detail below, as it was also used in Experiments 6 and 7); since they completed
449 this task quite quickly, for the remaining time they engaged in a conversation with the experimenter
450 about their studies. Participants in Experiments 2 and 3 completed the study online and, during the
451 break, they were asked to watch a short video from an episode of Peppa Pig and answer four
452 comprehension questions (to ensure they were paying attention).

453 Immediately after the break, all participants completed 8 trials in the retention phase (bottom
454 of Figure 1). On each retention trial, they again tapped on the picture of the cartoon character Peppa
455 Pig (top of the screen) and then heard an instruction to select the object corresponding to one of the
456 novel words (e.g., *Tap the cheem!*), while they observed three randomly-ordered pictures at the
457 bottom of the screen: the unfamiliar target object (the one that had appeared on the learning trial the
458 novel word was used on) and two other unfamiliar objects, which served as distractors. Of these,
459 one was a target object from a different trial, while the other had been also encountered by
460 participants in the learning phase, but on a filler trial, and had therefore not been named (see
461 Additional analyses in the *analysis_scripts* folder on the OSF, section 5, for a breakdown of
462 participants' errors by distractor type). Across retention trials, each unfamiliar target object
463 appeared twice (once as target, once as distractor) and each unfamiliar filler object also appeared
464 twice (always as a distractor, but paired with two different target words). Participants never
465 received any feedback about the accuracy of their choices. When pairing target objects with
466 distractors, we made sure that the average pairwise dissimilarity of the three objects was
467 comparable across trials (Mean = 0.8173, SD = 0.098, range [0.6490, 0.9624]; ratings from Horst
468 and Hout, 2016).

PREDICTION ERROR AND WORD LEARNING

469 All spoken instructions were recorded by a female native speaker of Scottish English with
470 child-directed prosody. Target words were recorded separately and combined with the spoken
471 contexts online, so that we could fully randomize object-word pairings for each participant. Trial
472 order was also randomized separately for each participant and each phase of the experiment.
473 Participants first completed the learning phase for all items and then completed the retention phase
474 (i.e., learning and retention were fully blocked, with no interleaving) The task was custom-coded in
475 HTML and Javascript. An OSF link to the code is available upon request: Since some of the visual
476 stimuli are protected by copyright, we are unfortunately unable to make all materials publicly
477 available.

478 **Data Analysis and Results.**

479 **Data analysis.** We analyzed participants' choices on learning trials (i.e., choosing the novel vs.
480 familiar object) and their accuracy on retention trials as a function of Constraint. For the retention
481 trials, accuracy was coded in terms of whether participants were able to retain the pairing of the
482 novel label with the novel object, regardless of whether they had chosen the novel object or the
483 familiar distractor during the learning phase. Additional analyses of retention accuracy controlled
484 for the choice made on the corresponding learning trial (Choice-at-learning) and were followed up
485 with separate analyses of retention trials for which the novel object had been chosen (Novel) on the
486 corresponding learning trial, and retention trials for which the familiar object had been chosen
487 during learning (Familiar) to check how previous referential choices affected retention. Fixed
488 effects were contrast coded and centered.

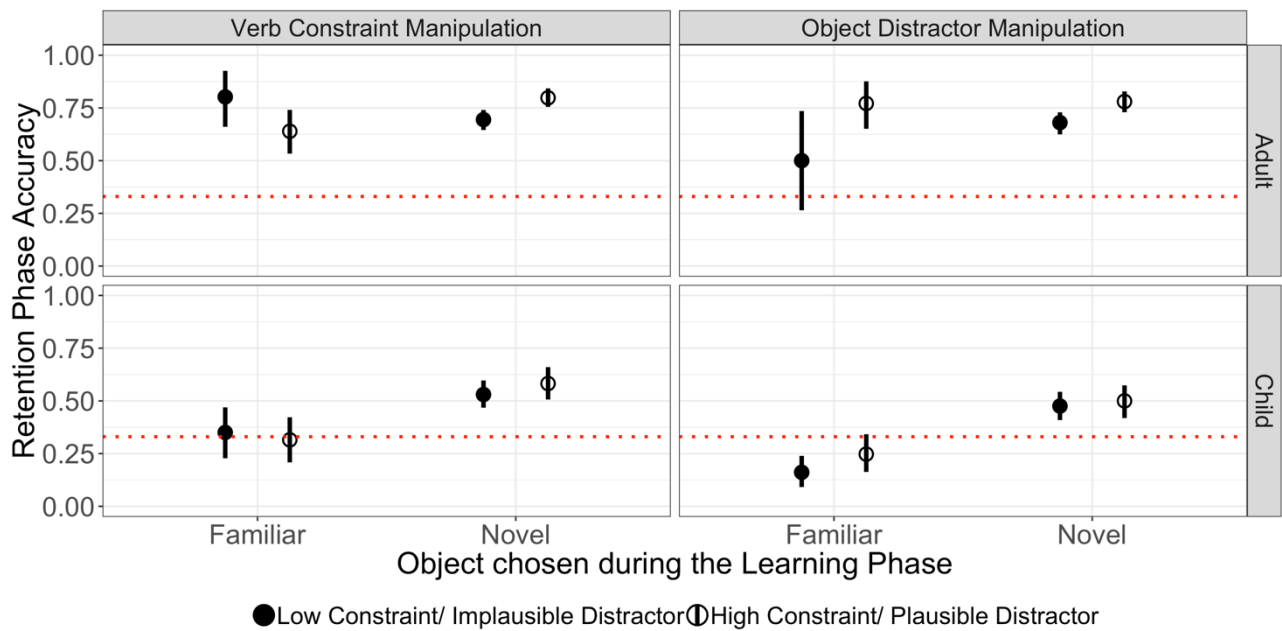
489 Since we combined data for three experiments, Experiment was added as an additional
490 factor with three levels and contrast coded; the first contrast compared performance in Experiment
491 1, which took place in the lab, to performance in the two online experiments (2 and 3), while the
492 second contrast compared performance in Experiment 2 to Experiment 3. The models included
493 interactions between these two contrasts and the fixed effect of interest (Constraint); for analyses of
494 retention accuracy, we initially also included interactions between the Experiment contrasts and

495 Choice-at-learning, but these more complex models did not converge. All analyses used generalized
496 linear mixed effects models with a logistic link function (function *glmer* from the *lme4* package;
497 Bates, Maechler, Bolker, & Walker, 2015) in R (R, Version 3.5.1). Random effects structure was
498 kept maximal, unless (1) correlations between random effects and/or (2) higher-order random
499 slopes had to be dropped to aid convergence (full model specifications available in the Analysis
500 Summary within the *analysis_scripts* folder, section 1, at the OSF link). Instead of *p* values, we
501 report 95% confidence intervals for model estimates from the *confint* function (method="Wald").
502 **Results.** To maximize power, we report a combined analysis of data from all three adult
503 experiments, but findings were highly consistent across all experiments (see Additional analyses in
504 the *analysis_scripts* folder, section 1, on the OSF for separate analyses for Experiments 1, 2 and 3),
505 and there were no significant differences between Experiments (either as main effects or
506 interactions with Constraint) in any of the analyses reported below (see Analysis Summary in the
507 *analysis_scripts* folder, section 1, on the OSF). Importantly, the planned replication (Experiment 3)
508 was successful ($z = 1.65$). Descriptive statistics for these and subsequent experiments are provided
509 in Table 2.

510 Accuracy on filler trials was 100%. During learning, adults were more likely to (correctly)
511 select the novel object on low constraint (92%) than high constraint trials (81%); this difference was
512 significant: log-odds $B = -1.59$, $SE = 0.26$, $z = -6.23$, $CI = [-2.53, -1.09]$. Conversely, on retention
513 trials, adults were more accurate for novel word-object pairs they had encountered on High
514 Constraint trials during the learning phase (76%) than on those they had encountered on Low
515 Constraint trials (69%); log-odds $B = 0.39$, $SE = 0.15$, $z = 2.65$, $CI = [0.10, 0.67]$; see Figure 2, top
516 left. This pattern was qualified by an interaction between Constraint and Choice-at-learning (log-
517 odds $B = 1.38$, $SE = 0.49$, $z = 2.81$, $CI = [0.42, 2.33]$), which indicated that it was driven by novel
518 (i.e., "correct") learning trials; log-odds $B = 0.66$, $SE = 0.17$, $z = 3.98$, $CI = [0.34, 0.98]$. In contrast,
519 retention of familiar (i.e., "inaccurate") learning trials tended to be worse for High Constraint items,
520 but this pattern was not reliable; $CI = [-1.68, 0.07]$.

PREDICTION ERROR AND WORD LEARNING

521 Figure 2. Retention accuracy (%) as a function of Verb Constraint (left) or type of Object Distractor
 522 (right) and of the referent chosen during learning (Familiar vs. Novel). The top panels report data
 523 from the adult experiments (Verb Constraint: Experiments 1-3; Object Distractor: Experiments 4-
 524 5), while the bottom panels report the child data (Verb Constraint: Experiment 6; Object Distractor:
 525 Experiment 7). Conditions where weaker expectations were violated are represented by a filled
 526 circle, while conditions where stronger expectations were violated are represented by an empty
 527 circle. The error bars represent 95% bootstrap CI's (1000 samples) over subjects. The dashed
 528 horizontal lines represent chance performance (33%).



529

530

531

532

533

534

535

536

537

538

539 Table 2. Descriptive statistics for all experiments.

Experiment		1-3	4-5	6	7
% Filler accuracy		100	>99	>99	96
(learning phase)					
% Novel object choices	High Constraint/ Plausible	81	86	66	63
(learning phase)	Distractor trials				
	Low Constraint/Implausible	92	96	77	69
	Distractor trials				
% Retention accuracy	High Constraint/ Plausible	80	79	57	50
(novel trials)	Distractor trials				
	Low Constraint/Implausible	69	68	52	49
	Distractor trials				
% Retention accuracy	High Constraint/ Plausible	61	73	30	25
(familiar trials)	Distractor trials				
	Low Constraint/Implausible	78	48	35	20
	Distractor trials				

540

541

542 **Discussion.**

543 In accord with prediction error-based theories of memory (Henson & Gagnepain, 2010),
 544 adults were more likely to retain a newly formed association between a word and its referent when
 545 that association disconfirmed a stronger expectation compared to a weaker one. Importantly, this is
 546 not merely a novelty effect: Pseudowords and unfamiliar objects were equally novel for participants
 547 across High and Low Constraint contexts. Critically, what changed was the strength of the prior
 548 expectations generated by the verbs.

549 However, adults were also much more likely to disregard mutual exclusivity when the
 550 constraint was High rather than Low (e.g., picking a picture of an apple as the referent for *cheem*

PREDICTION ERROR AND WORD LEARNING

551 more often after *eat* than *get*). This may suggest that, when contextual support is strong, adult word
552 learners may be more likely to infer that a novel word is a synonym for a highly expected familiar
553 word (e.g., *cheem* is a synonym for *apple*, or perhaps a type of apple). While this finding is
554 interesting in itself, and in line with previous evidence about adults' learning of novel word
555 meanings from context (Borovsky et al., 2010), it also means that we may have underestimated the
556 benefit of disconfirming strong expectations: Since familiar target objects were a much better fit
557 than unfamiliar objects after High Constraint contexts, adult learners may have found it more
558 difficult to revise their expectations following such contexts. Thus, we devised a second version of
559 the task where new unfamiliar objects were selected to better fit the High Constraint verbs.

560 Importantly, the new version also addressed a potential confound. Given that High and Low
561 Constraint conditions used different verbs, and that constraining verbs tend to be semantically
562 richer, it is possible that adult learners performed better in the High Constraint condition simply
563 because they could build richer and more distinctive representations for the word meanings,
564 providing more cues for retrieving information from memory. In the new version we therefore kept
565 sentential contexts constant and manipulated expectations by varying the plausibility of the familiar
566 object distractor instead.

567

568

569

570

571

572

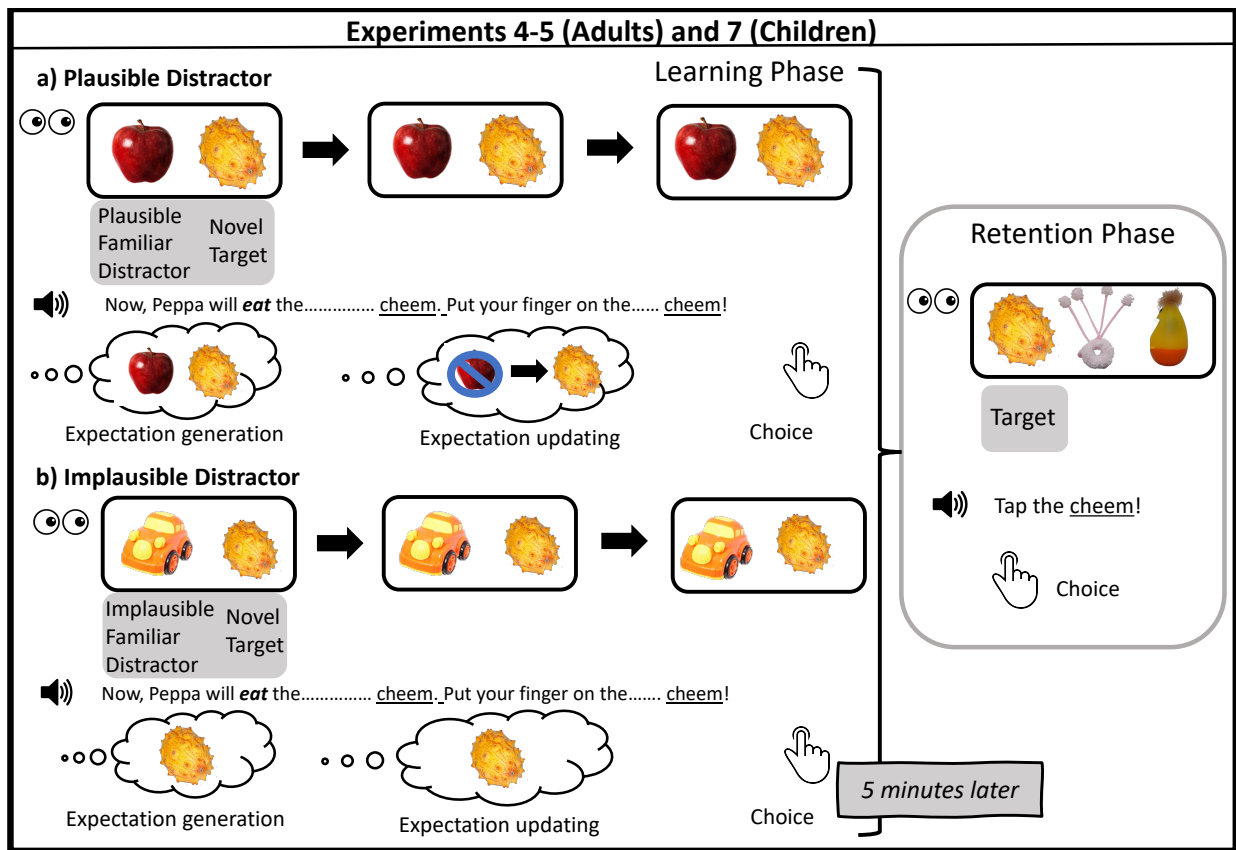
573

574

575

PREDICTION ERROR AND WORD LEARNING

576 Figure 3. Schematic depiction of the experimental design and graphical illustration of the processes
 577 at play during different types of learning trials in Experiments 4-5 and 7; a) Plausible Distractor
 578 trials, b) Implausible Distractor trials. Note that in (a) the strength of the expectation is larger for the
 579 plausible familiar distractor (apple) than the novel object (exotic fruit), but there is some
 580 expectation for the latter to be named – this reflects the findings from our post-test: The novel
 581 objects used in Experiments 4-5 and 7 were less of a good fit for the constraining verbs compared to
 582 the familiar objects, but they were also a better fit compared to the novel objects used in
 583 Experiments 1-3 and 6 (cf. Figure 1). In (b) the expectation updating step confirms the expectation
 584 generated initially (i.e., that the novel object will be named). We depict only learning trials on
 585 which participants choose the novel object as the referent of the novel word.



586
 587 **Experiments 4 and 5: Generating prediction errors using plausible distractor objects in adults**

588 These experiments were closely modelled on Experiments 1-3 but with two key
 589 modifications. First, we replaced all unfamiliar objects with objects that, while still unfamiliar,
 590 would better fit constraining verbs. For example, the target for *eat* was now an exotic fruit (see

591 Figure 3; a full list of materials is available in *the materials&lists* folder on the OSF). An additional
592 20 adults (6 male, 19 to 57 years of age), who participated in a similar post-test to the one
593 mentioned above, rated the new unfamiliar objects as more likely to undergo the actions referred to
594 by the constraining verbs ($M=4.91$), compared to the unfamiliar objects used in Experiments 1-3 (M
595 $= 3.08$, $t(34.72) = 5.40$, $p < .001$).

596 Secondly, learners were exposed only to semantically rich verbs (the constraining verbs
597 from Experiments 1-3). Rather than manipulating expectations by varying the verb, we instead
598 paired the same constraining verb (e.g., *eat*) either with a familiar object that fit its constraint (e.g.,
599 apple, as in Experiments 1-3) or with a different familiar object (e.g., car), which was implausible
600 given the verb (see Figure 3b). Thus, if semantic richness was responsible for the memory boost we
601 observed previously, we should now find no difference in retention accuracy using this design.
602 However, if the memory boost was driven by disconfirmed expectations, then we should find better
603 retention accuracy for trials with plausible than implausible familiar object distractors.

604 Implausible distractors should facilitate mapping of the novel word onto the correct,
605 unfamiliar target, even before the novel word is heard, so they should make it less likely that
606 participants will have their expectations disconfirmed (see Figure 3b); in other words, on
607 implausible distractor trials both the sentence context and the mutual exclusivity constraint should
608 bias participants to map the novel word onto the unfamiliar object. In contrast, on plausible
609 distractor trials, participants should still generate a strong expectation that the plausible familiar
610 distractors will be named (just as on high constraint trials in Experiments 1-3); in addition, they
611 may generate a weaker expectation that the unfamiliar object will be named (as this also fits the
612 constraint of the verb, though not as well as the familiar object). In any case, the occurrence of the
613 novel word should disconfirm the stronger expectation for the familiar distractor to be named,
614 generating prediction error (see Figure 3a).

615 **Participants.**

PREDICTION ERROR AND WORD LEARNING

616 One hundred and twenty-three adult participants took part online. Fifty-eight of these were students
617 from the University of Edinburgh (14 male, age range: 18 to 22, one participant did not provide age
618 information), who took part in Experiment 4. The remaining 65 participants were students from
619 Cardiff University (11 male, age range: 19 to 22, two participants did not provide age information)
620 and they took part in Experiment 5; two participants only completed the learning phase, so analyses
621 of retention accuracy are based on a sample size of 63 participants. Across the two experiments,
622 eleven participants were native speakers of a language other than English (6 in Experiment 4, 5 in
623 Experiment 5).

624 **Methods.**

625 The procedure was identical to Experiment 3. The design and materials were similar except for the
626 modifications described above: New unfamiliar target objects were chosen that provided a better fit
627 to the constraining verbs, and only sentences with constraining verbs were used, as we instead
628 varied the identity of the familiar distractor object, which could either be a good fit for the verb
629 (e.g., *apple* for *eat*; Plausible Distractor) or not (e.g., *car* for *eat*; Implausible Distractor).

630 Experiment 4 and 5 were almost identical replications of each other, with only a minor
631 variation in the assignment of items to conditions across the two experimental lists. We used two
632 lists in order to counterbalance the assignment of items to conditions (Plausible vs. Implausible
633 Distractor). While analyzing Experiment 4 data, we noticed that for a subset of the items, adults
634 were particularly likely to select the incorrect (familiar) distractor as the referent for the novel word
635 in the Plausible Distractor condition (but not in the Implausible Distractor condition), and these
636 items happened to cluster together in the counterbalancing (i.e., they all appeared in the Plausible
637 Distractor condition in the same list). As a result, one list led to fewer novel object selections during
638 the learning phase on Plausible Distractor than Implausible Distractor trials (88% vs. 99% Novel
639 choices), while the other did not (98% vs. 96% Novel choices). Since we were concerned this may
640 affect the results, we re-distributed item versions across lists before running Experiment 5. Lists for
641 both experiments are available in the *materials&lists* folder on the OSF.

642 **Results and Discussion.**

643 Since the two experiments yielded comparable findings, here we report combined analyses to
644 maximize power. Again, Experiment (contrast coded) and its interactions with the predictors of
645 interest (Distractor and Choice-at-Learning) were added to all models, and again there were no
646 significant differences between experiments, and there were no interactions modulating any of the
647 effects reported below. For separate analyses for each experiment, see Additional analyses in the
648 *analysis_scripts* folder, section 2, on the OSF. Accuracy on filler trials was higher than 99%.
649 During learning, adults were more likely to (correctly) select the target when the familiar distractor
650 was implausible (96%) compared to when it was a good fit (86%), though this difference was only
651 marginal; log-odds $B = -1.84$, $SE = 1.10$, $z = -1.68$, $CI = [-3.99, 0.31]$.

652 Most importantly, adult learners performed better at retention when their expectations had
653 been disconfirmed during learning (78%) than when they had not (67%); log-odds $B = 0.61$, $SE =$
654 0.19 , $z = 3.43$, $CI = [0.26, 0.96]$ (and this pattern did not depend on their choice during learning; CI
655 $= [-2.22, 0.49]$); see Figure 2, top right. Thus, disconfirmed expectations can enhance memory for
656 novel words, and it is unlikely that the findings from Experiments 1-3 were only due to differences
657 in semantic richness between verbs.

658 **Experiments 6-7: Children**

659 Having established that adults' memory for novel word-object associations is boosted by larger
660 prediction errors, we tested whether children would show similar effects using both the original
661 design (i.e., manipulating verb constraint as in Experiments 1-3) and the modified design (i.e.,
662 manipulating the distractor object as in Experiments 4-5).

663 **Methods.**

664 **Participants.** A refined power calculation based on data from Experiment 1-3 (total $N = 138$)
665 suggested that we may have overestimated the size of the effect in adults. This refined power
666 analysis indicated that a sample size of $N=80$ would achieve 83% power, so we aimed to recruit at

PREDICTION ERROR AND WORD LEARNING

667 least 80 children per experiment. The final sample sizes were 80 in Experiment 6 and 86 in
668 Experiment 7.

669 We had originally planned to test 2- and 3-year-olds because this age range sits at the
670 intersection between research on mutual exclusivity (e.g., Horst & Samuelson, 2008) and on
671 linguistic prediction (Borovsky et al., 2012; Mani & Huettig, 2012), but a few 4-year-olds were
672 included (10 in Experiment 6, 13 in Experiment 7) due to recruitment constraints; additional
673 analyses including the child's age in months did not reveal any age-related differences (see
674 Additional analyses in the analysis-scripts folder, section 3, on the OSF), so below we report
675 analyses that collapse across all ages. Children in Experiment 6 ($M_{\text{age}} = 38$ months, range = 25-56
676 months; 45 males, 35 females) were recruited from nurseries in the Edinburgh area, Edinburgh Zoo,
677 a local library, and from a database of families interested in research; children in Experiment 7
678 ($M_{\text{age}} = 38$ months, range = 24-59 months; 43 males, 43 females) were recruited from nurseries in
679 and around Cardiff, Techniquet (a science museum in Cardiff), from a database of families
680 interested in research, or through personal contacts. Written informed consent was obtained from all
681 caregivers and verbal assent from all children. All participants were exposed to English as one of
682 their home languages or at nursery, and some were exposed to at least one additional language (15
683 in Experiment 6, 16 in Experiment 7). Children who grow up bilingual may follow the mutual
684 exclusivity principle to a lesser extent than monolingual children (Byers-Heinlein & Werker, 2009),
685 so we added language background as a covariate in preliminary analyses. Since no differences were
686 found in these preliminary analyses, below we report analyses collapsing across number of
687 languages; note that in Byers-Heinlein and Werker (2009) the largest differences were observed
688 between monolingual and trilingual children and there were only two trilingual children in our
689 sample.

690 **Procedure.** The procedure was as similar as possible to the adult one. Children completed the task
691 on a touch-screen tablet. Although they were allowed to pace the task for themselves, the
692 experimenter monitored them closely to make sure they were paying attention to the spoken

693 instructions and, in case they appeared distracted, encouraged them to listen to the instructions
 694 again. During the break between the learning phase and the retention phase of the experiment,
 695 children completed a series of three tapping games involving known cartoon characters (as in
 696 Experiment 1); in each game, their task was to find the character named by the experimenter and
 697 “turn it” into a green tick mark by tapping on it with their finger. Experiment 6 used the same lists
 698 as Experiments 1-3 and Experiment 7 used the same lists as Experiment 4.

699 **Results.**

700 Children’s accuracy on filler trials was high (Exp. 6: >99%, Exp. 7: 96%). Like adults,
 701 children were more likely to (correctly) select the novel object on Low Constraint than High
 702 Constraint trials; Exp. 6: 77% vs. 66%); log-odds $B = -0.59$, $SE = 0.25$, $z = -2.33$, $CI = [-1.08,-$
 703 $0.09]$. Numerically, they were also more likely to select the novel object when the familiar
 704 distractor was implausible; Exp. 7: 69% vs. 63%), but this difference was not reliable; $CI = [-$
 705 $0.67,0.10]$.

706 In contrast to the adult findings, children’s retention of the novel word-object mappings was
 707 unaffected by the expectations they had generated during learning (see Figure 2, bottom panels). In
 708 Experiment 6, they were as accurate for pairs they had encountered on High (48%) or Low
 709 Constraint (48%) trials; $CI = [-0.35,0.30]$. In Experiment 7, they were similarly accurate regardless
 710 of whether the familiar distractor fit the verb well (41%) or was implausible (40%); $CI = [-$
 711 $0.26,0.36]$. These findings held even when we restricted the analysis to items for which children had
 712 chosen the novel referent during the learning phase (Experiment 6: $CI = [-0.18,0.60]$, Experiment 7:
 713 $CI = [-0.32,0.45]$).

714 Retention accuracy was much higher when children had (correctly) selected the novel object
 715 during learning, than when they had not (Exp. 6: 55% vs. 32%; Exp. 7: 50% vs. 23%); Experiment
 716 6: log-odds $B = 1.10$, $SE = 0.27$, $z = 4.01$, $CI = [0.56,1.63]$; Experiment 7: log-odds $B = 1.33$, $SE =$
 717 0.21 , $z = 6.44$, $CI = [0.92,1.73]$. However, choice at learning did not interact with our
 718 manipulations.

PREDICTION ERROR AND WORD LEARNING

719 While we set our sample size for each study using power analyses, these were based on adult
720 data, which are likely less variable than children's. However, combined analyses of data from both
721 Experiment 6 and 7 found no evidence for an effect of expectation strength on retention accuracy
722 (log-odds $B = 0.10$, $SE = 0.13$, $z = 0.79$, $CI = [-0.15, 0.35]$), despite their increased power. There
723 was also no indication that performance improved within the age range tested (log-odds $B = -0.001$,
724 $SE = 0.007$, $z = -0.09$, $CI = [-0.015, 0.014]$), nor that the size of the expectation strength effect was
725 larger for older children (log-odds $B = 0.01$, $SE = 0.02$, $z = 0.99$, $CI = [-0.01, 0.04]$; see Additional
726 analyses in the *analysis_scripts* folder, section 3, on the OSF).

727 Follow-up analyses combining data from all 7 experiments showed that, overall, adults'
728 choices at learning were affected by the strength prior expectations more than children's (log-odds
729 $B = -0.94$, $SE = 0.23$, $z = -4.04$, $CI = [-1.40, -0.48]$). Importantly, these analyses also confirmed that
730 adults' retention performance was affected by the strength prior expectations more than children's
731 (log-odds $B = 0.39$, $SE = 0.18$, $z = 2.18$, $CI = [0.04, 0.75]$).

732

733 **Discussion**

734 Unlike for adults, prediction errors did not enhance children's memory for word-referent
735 associations. This was despite clear evidence that children can generate expectations based on the
736 constraint of verbs even at age 2 (Mani & Huettig, 2012; recall that all of our constraining verbs
737 were English translations of stimuli in Mani and Huettig's German study). Moreover, children
738 clearly demonstrated sensitivity to the constraint manipulation in Experiment 6: Like adults, they
739 were much more likely to disregard mutual exclusivity when constraint was High (i.e., picking the
740 apple as the referent more often after *eat* than *get*), though their choices at learning were less
741 sensitive than adults' to the strength of prior expectations. Finally, although children's memory
742 performance was (unsurprisingly) lower than adults', it was still above chance, which suggests that,
743 although the task was difficult, children still encoded significant amounts of information during the

744 learning phase. Thus, our results cannot be explained by a floor effect. They suggest that prediction
745 errors play a surprisingly small role in how children encode word meanings.

746 Moreover, as we discuss below, other aspects of these data may be informative for models
747 of children's word learning. Children were strongly affected by their choices during learning: In
748 fact, while their retention was above chance-level (33% in this task) for words they had (correctly)
749 mapped onto the novel referent during learning, it was at chance for words they had instead mapped
750 onto the familiar referent. This suggests that children had only tracked one potential word-referent
751 mapping during this task (Stevens et al., 2017; Trueswell, Medina, Hafri, & Gleitman, 2013). We
752 return to this point in the General Discussion. Finally, while Horst and Samuelson (2008) found no
753 evidence for retention in 24-month-olds, we showed that children aged between 2 and 4 years were
754 able to retain the new word-referent mappings at above-chance levels over at least a 5-minute
755 period. This could suggest that children's retention abilities improve dramatically during the second
756 year of life, but note another important difference between our design and Horst and Samuelson's:
757 We presented the novel words in informationally rich, high constraint sentential contexts (e.g. *...eat*
758 *the cheem*), which may have facilitated more robust encoding of the word-referent mappings,
759 whereas they used only low constraint contexts (e.g., *get the cheem!*).

760

761

General Discussion

762 Can a prediction-error mechanism explain how adults and children encode associations
763 between novel word forms and their meanings? The evidence around this important question is
764 surprisingly mixed and, despite considerable evidence that both adults and children can process
765 language predictively, the role of prediction in the creation of new linguistic representations
766 remains poorly understood. In the introduction, we argued that a key hypothesis of error-driven
767 accounts of memory formation is that the disconfirmation of expectations should enhance memory
768 for the unexpected information and, importantly, the more so the stronger the initial expectation. In
769 this study, we tested this prediction in both 2-to-4-year-olds (2 experiments, combined N = 166)

770 and young adults (5 experiments, combined N = 259). Below, we summarize our findings and then
771 discuss their implications for our understanding of the mechanisms that support word learning and
772 their development.

773 There are two key findings. First, young adults are more likely to remember a novel word-
774 object association that has disconfirmed a stronger, compared to a weaker, expectation. We
775 established this finding (Experiments 1 and 2), directly replicated it (Experiment 3), and showed
776 that it still held when we modulated expectation strength through visual rather than linguistic
777 context (Experiments 4 and 5). Second, and in contrast to the adult findings, 2-to-4-year-olds'
778 memory was not enhanced by violations of stronger, compared to weaker expectations
779 (Experiments 6 and 7). This was despite the fact that children clearly generated linguistic
780 expectations: These expectations were strong enough to affect their referential choices (i.e.,
781 choosing the familiar object more often when it was more expected). Moreover, these expectations
782 also had an indirect effect on memory: When children failed to revise during the learning phrase,
783 they retained nothing about the novel objects and associated labels for the test phase. But when
784 words were mapped to novel objects during learning, expectation strength did not affect children's
785 retention.

786

787 **Prediction error shapes the encoding of linguistic information in adult memory: Implications**
788 **for models of word learning.**

789 Our adult findings clearly show that linguistic expectations shape the encoding of the link
790 between novel words and their meanings in memory, and can thus be viewed as an extension of the
791 PIMMS framework for memory (Henson & Gagnepain, 2010; Greve et al., 2017) to *linguistic*
792 representations. Importantly, these findings also have far-reaching consequences for computational
793 models of word learning. Such models have implemented a variety of different mechanisms, from
794 associative (i.e., Hebbian) learning (e.g., Kachergis, Yu, & Shiffrin, 2012; McMurray, Horst, &
795 Samuelson, 2012; Yu, Smith, Klein, & Shiffrin, 2007) to Bayesian inference (e.g., Xu &

796 Tenenbaum, 2007; Frank, Goodman, & Tenenbaum, 2009), from hypothesis testing (e.g., Stevens et
797 al., 2017; Trueswell et al., 2013; Yu et al., 2007) to the application to semantic-interpretation rules
798 (e.g., Siskind, 1996). However, with a few exceptions (Plaut & Kello, 1999; Ramskar et al. 2010;
799 Grimmick et al., 2019; Stevens et al., 2017), such mechanisms have not included error-driven
800 learning.

801 An error-driven learning mechanism is one that updates the current state of the model based
802 on the discrepancy between expected and observed inputs. By doing so, it can account for the role
803 played by prior expectations in learning: In our study, generating a stronger, but incorrect, prior
804 expectation led to the creation of a stronger memory trace for the correct word picture-mapping
805 (once the initial expectation was revised), suggesting that the generation of incorrect expectations
806 may benefit word learning. It is useful to contrast this with associative (Hebbian) learning: In its
807 simplest form, an associative word learner tracks the co-occurrences between words and referents,
808 augmenting the strength of the association between a word and a referent every time they co-occur
809 (e.g., Yu et al., 2007). More sophisticated associative models include parameters that let the
810 strength of associations decay over time, and can also model attention – that is, the fact that not all
811 possible word-referent associations are processed and stored equally (e.g., Kachergis et al., 2012).

812 However, associative models cannot straightforwardly account for the fact that association
813 strength depends on prior expectations. Recall that Grimmick et al. (2019) recently showed that
814 training adults on one set of word-referent mappings in a cross-situational learning paradigm, and
815 then changing the mappings, led to better memory performance for the items that had been changed
816 (i.e., initially incorrect items) than for those that had not. We argued that Grimmick et al.’s finding
817 also supports the hypothesis that prediction error is implicated in adult word learning and, indeed, in
818 order to reproduce their human data, Grimmick et al. augmented an associative word learning
819 model (Kachergis et al., 2012) with a prediction-error mechanism; the associative model by itself
820 could not reproduce their finding. Similarly, our findings suggest that adult word learning makes
821 use of a prediction-error mechanism.

822 Our findings can also be explained in terms of McMurray et al.'s (2012) competition-based
823 model of word learning. This model assumes that potential referents for a heard word compete with
824 each other, and that this process of "in-the-moment" competition during linguistic processing can
825 affect long-term learning (i.e., leading to changes in the weights representing the strength of
826 associations between words and their referents). In our study, competition levels were likely higher
827 on high-constraint and plausible distractor trials (compared to low constraint and implausible
828 distractor trials, respectively), and thus the novel target referent had to reach a higher level of
829 activation in order to be selected. If this higher activation translates into stronger association
830 weights, McMurray et al.'s model could explain the higher memory performance displayed by
831 adults for items encountered on those trials.

832 Note that other types of models can also be augmented with prediction-error mechanisms.
833 Recent years have seen the emergence of so-called hypothesis-testing models of word learning
834 (e.g., Trueswell et al., 2013). In these models, when learners hear a novel word, they generate a
835 single hypothesis about its referent, rather than tracking all possible associations between the word
836 and every co-occurring referent. If this hypothesis is confirmed on the next encounter, the
837 hypothesized word-referent mapping is retained, but if it happens to be disconfirmed, then the
838 learner needs to start afresh, as they have not retained any information from previous encounters;
839 see Berens, Horst, and Bird (2018) for evidence supporting this model using fMRI activation
840 patterns in the hippocampus during cross-situational word learning.

841 While the original hypothesis-testing model (Trueswell et al., 2013) includes processes of
842 expectation generation and error computation, it does not incorporate a prediction-error mechanism
843 because its learning following a disconfirmed expectation is *not* proportional to the strength of that
844 expectation. However, a recent modification of the original model, called PURSUIT, augments it
845 with a prediction-error mechanism where the amount of learning is proportional to expectation
846 strength (Stevens et al., 2017). We suggest that our findings are more compatible with this
847 augmented model than with the original hypothesis-testing model.

848 While an error-based learning mechanism straightforwardly explains our findings, we note
849 that they could also be accommodated within a Bayesian framework. Expectation generation would
850 be akin to positing a prior probability distribution, and expectations would then be updated based on
851 how surprising the data are given the prior, to derive a posterior probability distribution. On a high
852 constraint trial, most of the prior probability mass is placed on the expectation that the familiar
853 object will be mentioned next, while on a low constraint trial, it is distributed more evenly between
854 the familiar and unfamiliar object (see Figure 1). Thus, the very same data (i.e., the occurrence of
855 the novel word) will lead to a larger updating on a high constraint than low constraint trials, because
856 the novel word increases the probability that the unfamiliar object will be mentioned. However,
857 existing Bayesian models of word learning (Xu & Tenenbaum, 2007; Frank, et al., 2009) do not
858 include memory parameters, so it is unclear how they would account for the finding that larger
859 updating leads to enhanced retention. In contrast, this finding highlights the importance of building
860 models of word learning that account for the nature of memory.

861 More speculatively, our findings may also help link computational models of word learning
862 with the cognitive neuroscience of word learning. A large body of evidence implicates the
863 hippocampus in the initial stages of word learning in adults (Davis & Gaskell, 2009; Tagarelli,
864 Shattuck, Turkeltaub, & Ullman, 2019; Berens et al., 2018). According to the complementary
865 systems account of word learning (Lindsay & Gaskell, 2010), the hippocampus supports rapid,
866 initial acquisition of novel words, whereas the neocortex is responsible for slower consolidation,
867 typically following periods of sleep (see McClelland, McNaughton, & O'Reilly, 1995 for detailed
868 theoretical arguments in support of the complementary systems account of learning and memory).
869 Strong evidence for this account comes from the inability of patients with hippocampal lesions to
870 learn new words (see Cooper, Greve, & Henson, 2019, for a recent review and discussion).

871 Interestingly, the hippocampus is sensitive to novelty and unexpected events (e.g., Kumaran
872 & Maguire, 2006), and it is thought to encode not just episodic memories but also predictions about
873 future outcomes (e.g., Shohamy & Adcock, 2010). Our finding that prediction errors affect word

874 learning in adults, therefore, is consistent with a key role for the hippocampus in this process. Given
875 we did not find evidence for a role of prediction error in children word learning, an interesting
876 question for future research is whether there are significant developmental changes in the reliance
877 of word learning processes on the hippocampus (similarly to what has been shown for other areas of
878 cognitive development; Qin et al., 2014). It may also be that children only show sensitivity to
879 prediction error after a period of sleep consolidation (which was not included in our study).

880

881 **Do the mechanisms of word learning change across the lifespan?**

882 A key contribution of our study is to highlight a potential developmental discontinuity in the
883 mechanisms of word learning: While adult memory for novel word-referent mappings was affected
884 by the strength of prior expectations, we found no evidence that 2-to-4-year-olds' memory was
885 similarly affected. This raises the possibility that the above-described models of word learning,
886 which have often been evaluated based on adult data, may not automatically generalize to explain
887 children's behavior.

888 Our study is not the first to highlight differences between adults' and children's word
889 learning mechanisms. We have already mentioned Fitneva and Christiansen's (2017) work,
890 showing that 4-year-olds learn more when their expectations are confirmed, but adults learn more
891 when they are exposed to a higher proportion of unexpected word-referent mappings. But while
892 their findings suggest that children should benefit from generating correct expectations (see also
893 Benitez & Saffran, 2012, 2018), we found no difference in memory performance between trials on
894 which generating an incorrect expectation was more likely (high constraint) and those in which it
895 was less likely (low constraint).

896 Can Fitneva and Christiansen's (2017) findings be reconciled with ours? Incorrect
897 expectations may both hinder selection of the correct novel referent and benefit memory for it, if it
898 is selected. However, when we excluded all trials on which children (incorrectly) selected the
899 familiar referent, we still found no evidence for a difference in children's memory performance

900 between strong and weak expectation trials. Thus, it seems more likely that children at this age are
901 yet to develop the mechanism that makes memory sensitive to expectation strength.

902 Further, our results seem incompatible with Ramscar, Dye, and Klein (2013a), who argued
903 that children's word learning is more likely than adults' to be driven by an error-based mechanism
904 (rather than *less* likely, as our findings suggest). They devised a word learning task in which a
905 learner driven only by prediction error (Rescorla & Wagner, 1972; Ramscar et al., 2010) would
906 behave differently from one who additionally makes use of explicit inferences (e.g., reasoning by
907 exclusion). Participants were first exposed to three novel objects and two novel words. Two of the
908 objects co-occurred with only one of the words each, while the third object co-occurred with both of
909 the words. When children were presented with a third novel word at test (here, *wug*), they were less
910 likely to select this third object as a referent for *wug*. While none of the objects had co-occurred
911 with *wug* during the learning phase, the object that had co-occurred with two other words was the
912 least predictive of *wug*, and so it should be the least preferred choice of an error-driven learner
913 (Ramscar et al., 2010). In contrast, adults were more likely to select the third object than either of
914 the other objects, suggesting that they were more likely to explicitly reason by exclusion, choosing
915 the third object because they had already mapped each of the other two objects onto the word it had
916 co-occurred with.

917 While Ramscar et al.'s (2013a) findings suggest that child word learners track co-occurrence
918 information across multiple encounters, we note that in our study co-occurrence information was
919 identical across high and low constraint conditions. Instead, what (likely) changed across conditions
920 was participants' processing of the novel word and novel target object; for example, by violating a
921 strong linguistic expectation, we may have prompted deeper processing of the novel word and
922 object, which in turn would have led to enhanced encoding of the association between word and
923 object in memory.

924 We thus suggest that by the age of 4 children may be capable of accumulating information
925 using an error-driven learning rule to track which words and referents co-occur and which do not

926 (in accordance with Ramskar et al., 2013a), and form expectations about future co-occurrences. But
927 at this age the violation of such prior expectations does not yet lead to deeper processing and
928 encoding of unexpected information in memory. In sum, different mechanisms, with different
929 developmental trajectories, may underly our ability to track regularities in the environment (i.e.,
930 statistical learning; Yu & Smith, 2007) and to focus attention and cognitive resources on the
931 encoding of unexpected events.

932

933 **The effect of linguistic prediction on children’s word learning.**

934 Our findings suggest two conclusions about how prediction affects children’s learning. The
935 first conclusion is that children’s predictions affect what children learn, by guiding their attention,
936 but the second conclusion is that these predictions do not affect the strength of children’s memory
937 representations. These two conclusions may seem to contradict one another, but we propose they
938 can be reconciled with one another, and with findings from previous work (Reuter et al., 2019) by
939 carefully distinguishing the mechanisms involved.

940 First, the predictive strength of the sentence contexts affected the inferences that children
941 made about the likely referent of the novel word: They were more likely to choose the familiar
942 object (thus disregarding the mutual exclusivity constraint) when the sentence context led them to
943 expect a reference to this object. In turn, choosing the familiar object as the referent led to chance
944 performance at test, suggesting that children’s attention was focused on the selected referent, to the
945 extent that little information about the unselected referent was retained – a finding which,
946 incidentally, replicates previous studies (Aravind et al., 2018; Woodard, Gleitman, & Trueswell,
947 2016; but see Yurovsky & Frank, 2015) and is consistent with hypothesis-testing models of word
948 learning (Trueswell et al., 2013). Importantly, however, during the learning phase children still
949 selected the novel referent at above-chance rates, even when doing so required them to abandon a
950 prior expectation, and when they did select the novel referent during learning, they then
951 demonstrated above-chance retention of the association between the novel word and this novel

952 referent during the test phase. Thus, children were capable of revising and updating their
953 expectations based on the mismatch between those and the auditory input (i.e., when a novel word
954 occurred instead of the expected familiar one) on the majority of trials.

955 Second, the predictive strength of the sentence contexts did not affect children’s likelihood
956 of retaining the association between the novel word and the novel object. Thus, while the ability to
957 revise disconfirmed expectations may guide children to discover new linguistic information (i.e.,
958 one aspect of learning), we suggest it is not a key driver of *retention* of this information. This
959 interpretation allows our data to be reconciled with Reuter et al.’s (2019) finding that children who
960 show a stronger predict-and-revise looking pattern are also better at word learning. Recall that in
961 their study children’s performance at test was no greater in the high than the low constraint
962 condition (in fact, it was greater in the latter than the former), so they also found no evidence that
963 stronger expectations were associated with enhanced memory, when disconfirmed. What they did
964 find was that children who engaged less in prediction-and-revision were less likely to remember
965 high-constraint words, which is actually in line with our findings: When children did not engage in
966 mutual exclusivity reasoning during learning, then they had poor memory at test. Thus, Reuter et
967 al.’s findings concur with ours in suggesting that prediction-and-revision skills help reference
968 resolution in children, but do not affect retention, so long as reference is resolved to the object that
969 is later tested for retention.

970 There is however one caveat to these conclusions that is worth considering. Children’s
971 choices during learning were less sensitive to the strength of prior expectations compared to adults’.
972 This could be in part because children’s choices are often noisier than adults’ (e.g., due to lapses in
973 attention). But it is also expected because studies that have compared predictive skills between
974 children and adults have typically found stronger effects of prediction in adults (e.g., Gambi et al.,
975 2016, 2018; Borovsky et al., 2012). Moreover, the strength of prediction effects increases
976 throughout the pre-school years (Gambi, Jindal, Sharpe, Pickering, & Rabagliati, in press). This

PREDICTION ERROR AND WORD LEARNING

977 raises the possibility that prediction did not affect children's retention because they did not generate
978 expectations that were strong or consistent enough (unlike adults).

979 However, we think that this possibility is unlikely, because children clearly generated quite
980 strong expectations. As noted above, when children did follow their expectations and chose the
981 familiar referent (which they were more likely to do than adults in the High Constraint/Plausible
982 Distractor conditions; compare 66% novel referent selections in Experiment 6 and 63% in
983 Experiment 7 for children with 81% in Experiments 1-3 and 86% in Experiments 4-5 for adults),
984 this choice had a large impact on their memory performance during the retention phase. Similarly,
985 there may be a worry that the lack of expectation strength effects on children's memory is down to
986 the task being too difficult for children of this age, but as noted above children's performance was
987 well above chance when they selected the novel object during learning, which they did on most
988 trials.

989 In sum, we argue that, despite the use of different tasks during the learning phase and
990 different measures of learning (looking-while-listening vs. referent selection), as well as a slightly
991 different age range (3-to-5 vs. 2-to-4-year-olds), Reuter et al.'s (2019) findings are consistent with
992 our own: Both studies suggest that children's predictions affect reference resolution but are unlikely
993 to drive retention of new word-meaning mappings. Therefore, we disagree with Reuter et al.'s
994 suggestion that their findings show that children's memory for novel word-object associations is
995 supported by a prediction-error mechanism. Instead, we suggest that children who exhibited a
996 stronger predict-and-revise pattern were better at word learning in their study because they were
997 faster at processing sentences, and their higher processing speed allowed them to learn following
998 high-constraint sentences even though these initially biased their attention towards the incorrect
999 referent.

1000

1001 **Why did disconfirmed expectations not boost memory in young children?**

PREDICTION ERROR AND WORD LEARNING

1002 If pre-school aged children can generate linguistic expectations, and revise such
1003 expectations “in-the-moment” when they are disconfirmed, then why does prediction error not
1004 affect encoding of novel linguistic information in young children’s memory? Below we discuss two
1005 possible answers to this question.

1006 One possibility is that the null effect follows from children’s lack of fluency at completing
1007 the task, which follows from a recent proposal that violations of expectations only influence
1008 memory once inhibitory control skills are well-developed (Brod, Breitwieser, Hasselhorn, & Bunge,
1009 2019). In our task, inhibitory control skills would be important for quickly suppressing the
1010 generated expectation once a novel word is encountered, allowing fluent mapping to the correct
1011 referent. Brod and colleagues (2019) have proposed that this use of inhibitory control is still not
1012 apparent even in late childhood: They found that violating expectations did not enhance memory for
1013 new declarative knowledge in children aged 9 to 12 years, but that it did enhance memory in adults
1014 (Brod, Hasselhorn, & Bunge, 2018).

1015 This could potentially explain children’s difficulty with our task. While our child
1016 participants were able to inhibit selection of the strongly expected familiar object on the majority of
1017 trials, it is likely that they took longer than adults to focus attention on the novel object, by which
1018 time, activation of the novel word in their working memory may have already started to decay, and
1019 this could have led to a weaker binding of the word-object association. In sum, perhaps children
1020 were not able to re-direct their attention quickly enough to benefit from the stronger encoding of
1021 information following a larger prediction error. If this is the case, then our findings indicate that
1022 children may have already developed an error-based learning mechanism, but their memory for
1023 novel word-referent mappings does not benefit from this mechanism (at least in our paradigm)
1024 because of delays in children’s development of attentional skills.

1025 Alternatively, children may show relative insensitivity to disconfirmed expectations because
1026 doing so is in fact adaptive for their learning. Since children’s linguistic knowledge is so limited,
1027 their linguistic input is likely to deliver more surprises more frequently (i.e., unexpected words), at

PREDICTION ERROR AND WORD LEARNING

1028 least compared to adults. Children may therefore be more likely than adults to “expect the
1029 unexpected” (i.e., placing a higher likelihood on the eventuality of encountering unexpected words).
1030 While this means that unexpected words may not leave a particularly strong trace in children’s
1031 memory, it also allows attentional resources to be distributed more evenly across many mildly
1032 surprising words. This idea is supported by evidence that children are indeed sensitive to the
1033 predictability of the environment. For example, the so-called Goldilocks effect shows that infants
1034 and young children prefer to attend to input that is of intermediate predictability, neither too
1035 predictable nor too unpredictable given their current knowledge about the environment (Kidd,
1036 Piantadosi, & Aslin, 2014), and children can also quickly learn to expect the unexpected when they
1037 have been exposed to a speaker that talks about very unlikely events (Yurovsky, Case, & Frank,
1038 2017).

1039 We do not know of any research that shows that children’s memory becomes more sensitive
1040 to unpredictable information as they become more knowledgeable about the environment, but if
1041 children’s word learning does indeed benefit from encountering unexpected information that
1042 violates “core knowledge” (Stahl & Feigenson, 2017), this may suggest that children’s memory is
1043 more sensitive to unpredictable information in domains that the child is more knowledgeable about
1044 (because core knowledge is acquired very early on). Similarly, children may be more sensitive to
1045 prediction error when there is a conflict between internally-generated expectations and strong
1046 external cues (e.g., unambiguous referential cues, such as an adult’s pointing) compared to
1047 situations where there is a conflict between internally-generated expectations and the child’s
1048 preferred interpretation of a novel stimulus, as in our task. In the latter case, the error signal may be
1049 too weak or noisy because it is based on the child’s own developing knowledge of language,
1050 whereas strong error signals from the environment may play a much more important role in shaping
1051 children’s error-based learning.

1052

1053 **Conclusion**

1054 In sum, we showed that prediction error drives the encoding of novel word-object
1055 associations in adult memory, as associations were encoded more strongly when they violated a
1056 stronger compared to a weaker prior expectation. However, we found no effect of disconfirming a
1057 stronger versus weaker prior expectation on children’s memory. The adult findings represent a clear
1058 demonstration that at least one of the mechanisms underlying adult word learning is based on the
1059 computation of prediction errors. Thus, they set an important constraint on models of adult word
1060 learning. The lack of a comparable effect of prediction error on children memory was not due to
1061 children’s inability to generate linguistic expectations, nor to an inability to revise them when they
1062 proved incorrect. Instead, we suggest that children are either too slow to inhibit disconfirmed
1063 expectations or that they do not prioritize the processing of unexpected information as much as
1064 adults, because the environment is overall more unpredictable to them. These findings thus
1065 highlight an important developmental discontinuity in the mechanisms that underlie prediction’s
1066 role in language learning.

1067

1068 **Acknowledgements**

1069 This research was partly supported by a Leverhulme Trust Research Project Grant (RPG-2014-253)
1070 to H.R. and M.J.P and an ESRC Future Research Leaders grant to H.R. (ES/L01064X/1). We thank
1071 all the participating children, families, and nurseries, as well as Techniquet, Edinburgh City
1072 Council Central Library, and Edinburgh Zoo. We are very grateful to Zoe Williams and Lucie
1073 Smith for help with creating the stimuli and collecting the data in Cardiff.

1074

References

1075

1076 Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of
1077 subsequent reference. *Cognition*, 73(3), 247-264.

1078

1079 Ameel, E., Malt, B., & Storms, G. (2008). Object naming and later lexical development: From baby bottle to
beer bottle. *Journal of Memory and Language*, 58(2), 262-285.

PREDICTION ERROR AND WORD LEARNING

- 1080 Aravind, A., de Villiers, J., Pace, A., Valentine, H., Golinkoff, R., Hirsh-Pasek, K., . . . Wilson, M. S. (2018).
1081 Fast mapping word meanings across trials: Young children forget all but their first guess. *Cognition*,
1082 *177*, 177-188.
- 1083 Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4.
1084 *Journal of Statistical Software*, *67*(1), 1-48. doi:10.18637/jss.v067.i01
- 1085 Benitez, V. L., & Saffran, J. R. (2018). Predictable events enhance word learning in toddlers. *Current*
1086 *Biology*, *28*(17), 2787-2793. e2784.
- 1087 Benitez, V. L., & Smith, L. B. (2012). Predictable locations aid early object name learning. *Cognition*,
1088 *125*(3), 339-352.
- 1089 Berens, S. C., Horst, J. S., & Bird, C. M. (2018). Cross-situational learning is supported by propose-but-
1090 verify hypothesis testing. *Current Biology*, *28*(7), 1132-1136. e1135.
- 1091 Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age
1092 is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of*
1093 *Experimental Child Psychology*, *112*(4), 417-436.
- 1094 Borovsky, A., Kutas, M., & Elman, J. (2010). Learning to use words: Event-related potentials index single-
1095 shot contextual word learning. *Cognition*, *116*(2), 289-296.
- 1096 Brod, G., Breitwieser, J., Hasselhorn, M., & Bunge, S. A. (2019). Being proven wrong elicits learning in
1097 children—but only in those with higher executive function skills. *Developmental science*, e12916.
- 1098 Brod, G., Hasselhorn, M., & Bunge, S. A. (2018). When generating a prediction boosts learning: The
1099 element of surprise. *Learning and Instruction*, *55*, 22-31.
- 1100 Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: infants' language experience
1101 influences the development of a word-learning heuristic. *Developmental science*, *12*(5), 815-823.
- 1102 Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Stanford University Papers and Reports on*
1103 *Child Language Development*, *15*, 17-29.
- 1104 Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234-272.
- 1105 Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science.
1106 *Behavioral and Brain Sciences*, *36*(2), 181-204.

PREDICTION ERROR AND WORD LEARNING

- 1107 Cooper, E., Greve, A., & Henson, R. N. (2019). Little evidence for Fast Mapping (FM) in adults: A review
1108 and discussion. *Cognitive Neuroscience*, *10*(4), 196-209.
- 1109 Davis, M. H., & Gaskell, M. G. (2009). A complementary systems account of word learning: neural and
1110 behavioural evidence. *Philosophical Transactions of the Royal Society B: Biological Sciences*,
1111 *364*(1536), 3773-3800.
- 1112 Dell, G. S., & Chang, F. (2014). The P-chain: Relating sentence production and its disorders to
1113 comprehension and acquisition. *Philosophical Transactions of the Royal Society B: Biological*
1114 *Sciences*, *369*(1634), 20120394.
- 1115 Den Ouden, H. E., Friston, K. J., Daw, N. D., McIntosh, A. R., & Stephan, K. E. (2008). A dual role for
1116 prediction error in associative learning. *Cerebral Cortex*, *19*(5), 1175-1185.
- 1117 Dikker, S., Rabagliati, H., Farmer, T. A., & Pylkkänen, L. (2010). Early occipital sensitivity to syntactic
1118 category is based on form typicality. *Psychological Science*, *21*(5), 629-634.
- 1119 Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, *14*(2), 179-211.
- 1120 Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., . . . Stiles, J. (1994). Variability in
1121 early communicative development. *Monographs of the society for research in child development*,
1122 *59*(5), 1-173.
- 1123 Fitneva, S. A., & Christiansen, M. H. (2011). Looking in the wrong direction correlates with more accurate
1124 word learning. *Cognitive Science*, *35*(2), 367-380.
- 1125 Fitneva, S. A., & Christiansen, M. H. (2017). Developmental changes in cross-situational word learning: The
1126 inverse effect of initial accuracy. *Cognitive Science*, *41*, 141-161.
- 1127 Frank, M. C., Goodman, N. D., & Tenenbaum, J. B. (2009). Using speakers' referential intentions to model
1128 early cross-situational word learning. *Psychological Science*, *20*(5), 578-585.
- 1129 Friston, K. J. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B:*
1130 *Biological Sciences*, *360*(1456), 815-836.
- 1131 Friston, K. J. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*(2),
1132 127-138.

PREDICTION ERROR AND WORD LEARNING

- 1133 Gambi, C., Gorrie, F., Pickering, M. J., & Rabagliati, H. (2018). The development of linguistic prediction:
1134 Predictions of sound and meaning in 2- to 5-year-olds. *Journal of Experimental Child Psychology*,
1135 *173*, 351-370. doi:<https://doi.org/10.1016/j.jecp.2018.04.012>
- 1136 Gambi, C., Jindal, P., Sharpe, S., Pickering, M. J., & Rabagliati, H. (in press). The relation between
1137 preschoolers' vocabulary development and their ability to predict and recognize words. *Child*
1138 *Development*.
- 1139 Gambi, C., Pickering, M. J., & Rabagliati, H. (2016). Beyond Associations: Sensitivity to structure in pre-
1140 schoolers' linguistic predictions. *Cognition*, *157*, 340-351.
- 1141 [dataset] Gambi, C., Pickering, M. J., & Rabagliati, H. Disconfirmed expectations and word learning. Open
1142 Science Framework, 2020, doi:10.17605/OSF.IO/ZVN6U.
- 1143 Grimmick, C., Gureckis, T. M., & Kachergis, G. (2019). *Evidence of error-driven cross-situational word*
1144 *learning*. Paper presented at the 41st Annual Conference of the Cognitive Science Society, Montreal,
1145 Canada.
- 1146 Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception.
1147 *Behavioral and Brain Sciences*, *27*(3), 377-442.
- 1148 Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, *87*(1), B23-B34.
- 1149 Havron, N., de Carvalho, A., Fiévet, A. C., & Christophe, A. (2019). Three-to Four-Year-Old Children
1150 Rapidly Adapt Their Predictions and Use Them to Learn Novel Word Meanings. *Child*
1151 *Development*, *90*(1), 82-90.
- 1152 Henderson, L. M., Weighall, A. R., Brown, H., & Gareth Gaskell, M. (2012). Consolidation of vocabulary is
1153 associated with sleep in children. *Developmental Science*, *15*(5), 674-687.
- 1154 Henson, R. N., & Gagnepain, P. (2010). Predictive, interactive multiple memory systems. *Hippocampus*,
1155 *20*(11), 1315-1326.
- 1156 Horst, J. S., & Hout, M. C. (2016). The Novel Object and Unusual Name (NOUN) Database: A collection of
1157 novel images for use in experimental research. *Behavior Research Methods*, *48*(4), 1393-1409.
- 1158 Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*,
1159 *13*(2), 128-157.

PREDICTION ERROR AND WORD LEARNING

- 1160 Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research, 1626*,
1161 118-135. doi:10.1016/j.brainres.2015.02.014
- 1162 Hulme, R. C., Barsky, D., & Rodd, J. M. (2019). Incidental Learning and Long-Term Retention of New
1163 Word Meanings From Stories: The Effect of Number of Exposures. *Language Learning, 69*(1), 18-
1164 43.
- 1165 Kachergis, G., Yu, C., & Shiffrin, R. M. (2012). An associative model of adaptive inference for learning
1166 word-referent mappings. *Psychonomic Bulletin & Review, 19*(2), 317-324.
- 1167 Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks effect in infant auditory attention. *Child*
1168 *Development, 85*(5), 1795-1804.
- 1169 Kumaran, D., & Maguire, E. A. (2006). An unexpected sequence of events: mismatch detection in the human
1170 hippocampus. *PLoS Biology, 4*(12), e424.
- 1171 Kuperberg, G. R., & Jaeger, T. (2016). What do we mean by prediction in language comprehension?
1172 *Language, Cognition and Neuroscience, 31*(1), 32-59.
- 1173 Lindsay, L., Gambi, C., & Rabagliati, H. (2019). Preschoolers optimize the timing of their conversational
1174 turns through flexible coordination of language comprehension and production. *Psychological*
1175 *Science, 30*(4), 504-515.
- 1176 Lindsay, S., & Gaskell, M. G. (2010). A complementary systems account of word learning in L1 and L2.
1177 *Language Learning, 60*, 45-63.
- 1178 Lukyanenko, C., & Fisher, C. (2016). Where are the cookies? Two- and three-year-olds use number-marked
1179 verbs to anticipate upcoming nouns. *Cognition, 146*, 349-370.
1180 doi:<http://dx.doi.org/10.1016/j.cognition.2015.10.012>
- 1181 Mani, N., Daum, M. M., & Huettig, F. (2016). "Pro-active" in many ways: Developmental evidence for a
1182 dynamic pluralistic approach to prediction. *Quarterly Journal of Experimental Psychology, 69*(11),
1183 2189-2201.
- 1184 Mani, N., & Huettig, F. (2012). Prediction during language processing is a piece of cake - but only for
1185 skilled producers. *Journal of Experimental Psychology: Human Perception and Performance, 38*(4),
1186 843-847.

PREDICTION ERROR AND WORD LEARNING

- 1187 Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children.
1188 *Nature*, 385(6619), 813-815.
- 1189 McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning
1190 systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist
1191 models of learning and memory. *Psychological Review*, 102(3), 419-457.
- 1192 McMurray, B., Horst, J. S., & Samuelson, L. K. (2012). Word learning emerges from the interaction of
1193 online referent selection and slow associative learning. *Psychological Review*, 119(4), 831.
- 1194 Niv, Y., & Schoenbaum, G. (2008). Dialogues on prediction errors. *Trends in Cognitive Sciences*, 12(7),
1195 265-272.
- 1196 Pickering, M. J., & Gambi, C. (2018). Predicting while comprehending language: A theory and review.
1197 *Psychological Bulletin*, 144(10), 1002.
- 1198 Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension.
1199 *Behavioral and Brain Sciences*, 36(4), 329-392. doi:<http://dx.doi.org/10.1017/S0140525X12001495>
- 1200 Plaut, D. C. and Kello, C. T. (1999). The emergence of phonology from the interplay of speech
1201 comprehension and production: A distributed connectionist approach. In B. MacWhinney (Ed.), *The*
1202 *emergence of language* (pp. 381-415). Mahwah, NJ: Erlbaum.
- 1203 Qin, S., Cho, S., Chen, T., Rosenberg-Lee, M., Geary, D. C., & Menon, V. (2014). Hippocampal-neocortical
1204 functional reorganization underlies children's cognitive development. *Nature Neuroscience*, 17(9),
1205 1263-1269.
- 1206 R. (Version 3.5.1) [Computer Software]. Vienna, Austria: R Development Core Team. Retrieved from
1207 <http://www.R-project.org>
- 1208 Rabagliati, H., Gambi, C., & Pickering, M. J. (2016). Learning to predict or predicting to learn? *Language,*
1209 *Cognition, and Neuroscience*, 31(1), 94-105. doi:10.1080/23273798.2015.1077979
- 1210 Ramscar, M., Dye, M., & Klein, J. (2013a). Children value informativity over logic in word learning.
1211 *Psychological Science*, 24(6), 1017-1023.
- 1212 Ramscar, M., Dye, M., & McCauley, S. M. (2013b). Error and expectation in language learning: The curious
1213 absence of "mouses" in adult speech. *Language*, 89(4), 760-793.

PREDICTION ERROR AND WORD LEARNING

- 1214 Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The effects of feature-label-order and
1215 their implications for symbolic learning. *Cognitive science*, 34(6), 909-957.
- 1216 Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness
1217 of reinforcement and nonreinforcement. *Classical conditioning II: Current research and theory*, 2,
1218 64-99.
- 1219 Reuter, T., Borovsky, A., & Lew-Williams, C. (2019). Predict and redirect: Prediction errors support
1220 children's word learning. *Developmental psychology*, 55(8), 1656-1665.
- 1221 Rommers, J., & Federmeier, K. D. (2018). Lingered expectations: A pseudo-repetition effect for words
1222 previously expected but not presented. *NeuroImage*, 183, 263-272.
- 1223 Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating
1224 errors. *Nature*, 323(6088), 533-536.
- 1225 Samuelson, L. K., & McMurray, B. (2017). What does it take to learn a word?. *Wiley Interdisciplinary*
1226 *Reviews: Cognitive Science*, 8(1-2), e1421.
- 1227 Shohamy, D., & Adcock, R. A. (2010). Dopamine and adaptive memory. *Trends in Cognitive Sciences*,
1228 14(10), 464-472.
- 1229 Siskind, J. M. (1996). A computational study of cross-situational techniques for learning word-to-meaning
1230 mappings. *Cognition*, 61(1-2), 39-91.
- 1231 Spiegel, C., & Halberda, J. (2011). Rapid fast-mapping abilities in 2-year-olds. *Journal of Experimental*
1232 *Child Psychology*, 109(1), 132-140.
- 1233 Stahl, A. E., & Feigenson, L. (2017). Expectancy violations promote learning in young children. *Cognition*,
1234 163, 1-14.
- 1235 Stevens, J. S., Gleitman, L. R., Trueswell, J. C., & Yang, C. (2017). The pursuit of word meanings.
1236 *Cognitive Science*, 41, 638-676.
- 1237 St. John, M. F., & McClelland, J. L. (1990). Learning and applying contextual constraints in sentence
1238 comprehension. *Artificial intelligence*, 46(1-2), 217-257.
- 1239 Tagarelli, K. M., Shattuck, K. F., Turkeltaub, P. E., & Ullman, M. T. (2019). Language learning in the adult
1240 brain: A neuroanatomical meta-analysis of lexical and grammatical learning. *NeuroImage*, 193, 178-
1241 200.

PREDICTION ERROR AND WORD LEARNING

- 1242 Trueswell, J. C., Medina, T. N., Hafri, A., & Gleitman, L. R. (2013). Propose but verify: Fast mapping meets
1243 cross-situational word learning. *Cognitive Psychology*, *66*(1), 126-156.
- 1244 Vlach, H. A. (2019). Learning to Remember Words: Memory Constraints as Double-Edged Sword
1245 Mechanisms of Language Development. *Child Development Perspectives*, *13*(3), 159-165.
- 1246 Vlach, H. A., & DeBrock, C. A. (2017). Remember dax? Relations between children's cross-situational word
1247 learning, memory, and language abilities. *Journal of Memory and Language*, *93*, 217-230.
- 1248 Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word
1249 learning. *Journal of Experimental Child Psychology*, *108*(2), 394-401.
- 1250 Vlach, H., & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children's
1251 retention of learned words. *Frontiers in psychology*, *3*, doi:10.3389/fpsyg.2012.00046.
- 1252 Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related
1253 brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish
1254 sentence reading. *Journal of Cognitive Neuroscience*, *16*(7), 1272-1288.
- 1255 Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology*, *11*(18), R729-R732.
- 1256 Woodard, K., Gleitman, L. R., & Trueswell, J. C. (2016). Two-and three-year-olds track a single meaning
1257 during word learning: Evidence for Propose-but-verify. *Language Learning and Development*,
1258 *12*(3), 252-261.
- 1259 Ylinen, S., Nora, A., Leminen, A., Hakala, T., Huotilainen, M., Shtyrov, Y., & Mäkelä, J. P. (2014). Two
1260 distinct auditory-motor circuits for monitoring speech production as revealed by content-specific
1261 suppression of auditory cortex. *Cerebral Cortex*, bht351.
- 1262 Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics.
1263 *Psychological Science*, *18*(5), 414-420.
- 1264 Yu, C., Smith, L. B., Klein, K. A., & Shiffrin, R. M. (2007). *Hypothesis testing and associative learning in*
1265 *cross-situational word learning: Are they one and the same?* Paper presented at the Proceedings of
1266 the Annual Meeting of the Cognitive Science Society.
- 1267 Yurovsky, D., Case, S., & Frank, M. C. (2017). Preschoolers flexibly adapt to linguistic input in a noisy
1268 channel. *Psychological Science*, *28*(1), 132-140.

PREDICTION ERROR AND WORD LEARNING

- 1269 Yurovsky, D., & Frank, M. C. (2015). An integrative account of constraints on cross-situational learning.
1270 *Cognition*, 145, 53-62.
- 1271 Yurovsky, D., & Frank, M. C. (2017). Beyond naïve cue combination: Salience and social cues in early word
1272 learning. *Developmental Science*, 20(2), e12349.
- 1273 Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological review*, 114(2),
1274 245-272.
- 1275