Received: 13 March 2023 | Revised: 19 December 2023 | Accepted: 26 February 2024

DOI: 10.1111/desc.13498

Developmental Science WILEY

RESEARCH ARTICLE

Children extract a new linguistic rule more quickly than adults

Sarah Berger | Laura J. Batterink

Department of Psychology, University of Western Ontario, London, Canada

Correspondence
Laura J. Batterink, Department of Psychology, University of Western Ontario, London, Canada.
Email: lbatter@uwo.ca

Funding information
Natural Sciences and Engineering Research Council of Canada

Abstract

Children achieve better long-term language outcomes than adults. However, it remains unclear whether children actually learn language *more quickly* than adults during real-time exposure to input—indicative of true superior language learning abilities—or whether this advantage stems from other factors. To examine this issue, we compared the rate at which children (8–10 years) and adults extracted a novel, hidden linguistic rule, in which novel articles probabilistically predicted the animacy of associated nouns (e.g., "gi lion"). Participants categorized these two-word phrases according to a second, explicitly instructed rule over two sessions, separated by an overnight delay. Both children and adults successfully learned the hidden animacy rule through mere exposure to the phrases, showing slower response times and decreased accuracy to occasional phrases that violated the rule. Critically, sensitivity to the hidden rule emerged much more quickly in children than adults; children showed a processing cost for violation trials from very early on in learning, whereas adults did not show reliable sensitivity to the rule until the second session. Children also showed superior generalization of the hidden animacy rule when asked to classify nonword trials (e.g., "gi badupi") according to the hidden animacy rule. Children and adults showed similar retention of the hidden rule over the delay period. These results provide insight into the nature of the critical period for language, suggesting that children have a true advantage over adults in the rate of implicit language learning. Relative to adults, children more rapidly extract hidden linguistic structures during real-time language exposure.

KEYWORDS
critical period, generalization, grammar learning, implicit learning, language learning, rule learning

Research Highlights

- Children and adults both succeeded in implicitly learning a novel, uninstructed linguistic rule, based solely on exposure to input.
- Children learned the novel linguistic rules much more quickly than adults.
- Children showed better generalization performance than adults when asked to apply the novel rule to nonsense words without semantic content.
- Results provide insight into the nature of critical period effects in language, indicating that children have an advantage over adults in real-time language learning.

1 | INTRODUCTION

A longstanding idea is that children are better equipped for language learning than adults (Johnson & Newport, 1989; Lenneberg et al., 1967; Long, 1990; Weber-Fox & Neville, 1996). Evidence for this idea comes from studies showing that younger learners achieve higher ultimate proficiency, both for a second language (Birdsong & Molis, 2001; DeKeyser, 2000; Hartshorne et al., 2018; Johnson & Newport, 1989; Weber-Fox & Neville, 1996) as well for a first (sign) language (Mayberry, 1993; Newport, 1990). Importantly, however, not all aspects of language are equally impacted by delays in acquisition (DeKeyser, 2000; Johnson & Newport, 1989; Newport, 1990; Newport et al., 2001; Weber-Fox & Neville, 1996). Vocabulary, semantic processing, and more salient aspects of morphosyntax, such as basic word order, are generally learned well even by late learners. In contrast, phonology and more complex and/or less salient aspects of morphology and grammar are much more impacted by delays in acquisition.

1.1 | Are children better at implicit language learning?

What gives children this long-term advantage in attaining structural aspects of language? A commonly proposed explanation is that children use different learning mechanisms for language, with children using more implicit language learning mechanisms and adults using more explicit learning mechanisms (DeKeyser, 2000; Elman, 1993; Kareev, 1995; Lenneberg et al., 1967; Newport, 1990; Paradis, 2009; Ullman, 2001). Here, we define implicit learning as a learning process that proceeds through exposure to positive examples, in the absence of feedback, explicit instruction, or intention to learn, and without necessitating conscious awareness of what has been learned (e.g., Frensch & Rünger, 2003; Nissen & Bullemer, 1987; Perruchet & Pacton, 2006; Reber, 1989)[1]. In particular, many theories suggest that adults' better cognitive abilities—including long-term memory, working memory, attention and cognitive control—may paradoxically make language learning more difficult. Due to competitive interactions between implicit and explicit learning mechanisms (e.g., Filoteo et al., 2010; Fletcher et al., 2005; Foerde et al., 2006; Nemeth et al., 2013; Poldrack et al., 2001; Smalle et al., 2022), adults' mature higher-level cognitive abilities may lead them to favor explicit, conscious strategies over implicit mechanisms when faced with the task of learning a new language. In contrast, younger learners—with their smaller cognitive capacities—are thought to rely predominantly on implicit or procedural learning mechanisms, which eventually produces a language attainment advantage over older learners (e.g., Newport, 1990; Paradis, 2009; Ullman, 2001). By these accounts, children are thought to have an acquisition advantage primarily for elements of language that are hard to learn explicitly, such as sequential components of language that involve extracting reoccurring units from input. These include a language's more complex and/or less prominent syntactic, morphological, and phonological rules (DeKeyser, 2000).

One such prominent model is Newport's "Less-is-More" account (Newport, 1988, 1990), which suggests that children's reduced working memory capacity leads them to perceive and store smaller component parts of complex linguistic stimuli. In turn, this allows children to extract the core components of language, such as small morphological units. In contrast, adults with their larger cognitive capacities may be biased towards storing whole word-meaning mappings, at the expense of detecting and extracting smaller relevant morphemes. A related account (Ramscar & Gitcho, 2007; Thompson-Schill et al., 2009) suggests that children's more limited cognitive control abilities leads them to learn in a more "unsupervised" manner, rather than selectively attending to and controlling their learning as adults do. This results in children learning the most consistent, frequent patterns in the input, which in turn facilitates grammar learning and other aspects of linguistic generalization. Yet another neurobiological theory suggests that children's language learning advantage can be traced to a developmental shift in reliance from the brain's fast-maturing procedural memory system to its later-maturing declarative memory system (Ullman, 2001, 2004, 2005). Aspects of grammar that are supported by the procedural memory system in early development may come to be acquired by the more dominant declarative memory system later in development, causing older learners to rely on effortful, conscious application of rules when using grammatical forms.

Parallel findings outside the domain of language also suggest that, relative to adults, children's poorer selective attention leads them to more deeply process and learn about task-extraneous information (Blanco & Sloutsky, 2019; Decker et al., 2015; Jung et al., 2023; Plebanek & Sloutsky, 2017). In turn, this may allow children to more effectively capture environmental contingencies that are not necessarily goal-relevant, oftentimes producing more robust incidental learning. For example, after completing a visual search task on arrays of artificial creatures, children's memory for the creatures with search-irrelevant features was better than adults', suggesting that the children had distributed their attention across both relevant and irrelevant information (Plebanek & Sloutsky, 2017). Although these various theories and findings differ in many respects, they all point to the idea that children, due to their immature high-level cognitive capacities, may rely more on implicit learning mechanisms when processing input, which in turn may give them an advantage when it comes to extracting hidden patterns from language.

Compelling evidence for this prediction would come from showing that children learn abstract, uninstructed, incidental regularities in language input more efficiently than adults—that is, that they have a faster short-term *learning rate* for linguistic patterns (conceptually distinct from long-term ultimate attainment; cf. Hartshorne et al., 2018; Krashen et al., 1979). However, very few studies have directly tested this prediction. Although there is work going back decades that has compared children and adults' short-term language learning under equivalent conditions (e.g., Asher & Price, 1967; Ferman & Karni, 2010; Hudson Kam & Newport, 2009; Krashen et al., 1979; Lichtman, 2016; Long, 1990; Snow & Hoefnagel-Höhle, 1978), most of these studies have used testing methods in which performance is likely to benefit significantly from explicit knowledge, and which may thereby potentially (inadvertently) favor older learners. Thus, perhaps not surprisingly, most findings in this literature suggest that younger children actually take *longer* to learn new grammar compared to older learners. For example, Ferman and Karni (2010) trained 8-year-olds, 12-year-olds and adults on an artificial morphological rule over 10 daily training sessions, with training consisting of a forced-choice grammaticality judgment tasks and a production task that required participants to generate the correct verb form. On these tasks, adults outperformed children, and older children outperformed younger children. However, as has been previously pointed out (Hartshorne et al., 2018; Long, 1990), these production-based and grammaticality judgment tasks are explicit and cognitively demanding in nature. Such tasks can potentially place older learners at an advantage, allowing them to benefit from conscious strategies, existing knowledge, and/or test-taking experience.

In contrast to most previous work, one recent study did find evidence that children show more rapid implicit language learning than adults (Smalle et al., 2017). This study specifically examined the learning of novel phonotactic constraints—the language-specific restrictions that govern permissible sound sequences in spoken language. Children (9–10 years) and adults were asked to rapidly recite syllable sequences that conformed to a novel phonotactic rule, with speech errors that were consistent with the novel constraint providing an index of implicit learning. Children showed reliable evidence of phonotactic learning very quickly, after exposure to only 24 sequences, whereas adults did not show evidence of learning until the second day, after a period of consolidation that contained sleep. These results may represent some of the strongest evidence to date that children have an underlying *rate* advantage for the implicit extraction of linguistic regularities. Notably, the implicit, performance-based measure of learning used in this study may have put children and adults on a more equal playing field, highlighting the importance of using undemanding, implicit and/or child-friendly measures in developmental studies of learning.

In light of these conflicting findings, one major goal of the current study was to clarify whether children show better implicit learning of hidden linguistic regularities, or whether their long-term advantage for language may instead by driven by other factors, such as environmental and social considerations (Flege & Liu, 2001; Flege et al., 1999; Hakuta et al., 2003; Marinova-Todd et al., 2000). In particular, it remains to be determined whether the advantage observed by Smalle and colleagues (2017) is specific to phonotactic learning, or also holds for other forms of implicit language learning, such as the acquisition of abstract grammar rules.

1.2 | Are there developmental differences in the long-term retention of linguistic regularities?

In addition to learning rate within a single training session, children's apparent superiority for language could also in principle be driven by better *long-term retention* of newly acquired linguistic representations, which would result in cumulative learning advantages over time. Supporting this possibility, some evidence suggests that children may retain recently learned verbal or episodic information better than adults (Bishop et al., 2012; Smalle et al., 2018; Wang et al., 2018). In one study, children and adults were asked to repeat novel wordforms (e.g., *preskrimskee*) in two sessions separated by 1 h (Bishop et al., 2012). Children's performance remained stable over the 1-h break whereas adults' performance declined, and this retention advantage in children was not attributable to differences in baseline learning. Similarly, another study using the Hebb repetition paradigm found that children showed better delayed retention of the implicit Hebb sequences relative to adults across multiple offline periods, ranging from 4 h up to 1 year later (Smalle et al., 2018). Again, children's retention advantage persisted even when controlling for baseline learning scores. Additional evidence from motor learning paradigms suggests that memory stabilization processes after learning may occur much more rapidly in children than adults (Adi-Japha et al., 2014; Ashtamker & Karni, 2013), and be less prone to subsequent interference (Dorfberger et al., 2007).

Sleep may further contribute to developmental differences in consolidation and retention, over and above any nonspecific effects of a delay period. In their review paper, Wilhelm, Metzkow-Mészàros, et al. (2012) conclude that children's longer and richer slow-wave sleep preferentially benefits declarative memories while preventing consolidation of procedural memories, a trade-off not generally seen in adults (Fischer et al., 2007; Giganti et al., 2014; Henderson et al., 2012; Peiffer et al., 2020; Prehn-Kristensen et al., 2009; Wilhelm et al., 2008). For example, in one study, 7–11-year-old children and adults completed a serial reaction time task, in which they pressed buttons corresponding to a hidden probabilistic pattern, and then either slept or stayed awake before completing the task again (Fischer et al., 2007). After sleep, adults showed a gain in implicit sequence knowledge, as reflected by facilitated response times to the pattern. In contrast, the reaction time (RT) difference to sequential versus control blocks was *reduced* after sleep in children, reflecting a reduction in implicit sequence knowledge. A subsequent study using similar methodology found that 8–11-year-old children demonstrated greater gains in explicit knowledge of the sequence compared to adults following a night of sleep, and that these gains were correlated with their slow-wave activity (Wilhelm et al., 2013). Altogether, these results suggest that in children, a consolidation period including sleep may preferentially strengthen explicit knowledge over implicit or procedural knowledge, and—when relevant—more strongly facilitate the transformation from implicit to explicit knowledge than in adults. By comparison, in adults, a delay period containing sleep seems to benefit the consolidation of implicit memories to a greater extent than in children.

With this background in mind, an additional aim of the current study was to better understand developmental differences in the retention of a hidden, abstract linguistic rule over a 12-h delay period containing sleep. We also sought to determine whether a consolidation period including sleep would differentially spur the transformation of implicit pattern knowledge to a more generalizable and/or consciously available rule in children compared to adults, as suggested by some prior work (Wilhelm et al., 2013).

1.3 | Current study

In the current study, we tested two interrelated hypotheses. Our first hypothesis was that children should become sensitive to a novel, hidden grammatical regularity more quickly than adults. A second, more exploratory hypothesis was that children, relative to adults, may show enhanced explicit, generalizable knowledge of the hidden linguistic rule after a 12-h period containing sleep. We tested children between 8 and 10 years old, as this age range is comparable to that of many prior studies of interest (e.g., Fischer et al., 2007; Smalle et al., 2017, 2018; Wilhelm et al., 2013) and also occurs well before the estimated cutoff point at which attainment of native-like proficiency becomes unlikely (age 12; Hartshorne et al., 2018).

To test these hypotheses, we created a child-friendly version of our previous implicit learning paradigm (Batterink et al., 2014), which involved exposing participants to a novel, hidden linguistic animacy rule. Specifically, both child and adult participants listened to two-word phrases that included one of four novel articles (gi, ro, ul, ne) followed by a subsequent English noun (e.g., *gi lion*). Participants were explicitly instructed that two of the novel words meant the accompanying noun was near, and the other two meant it was far. However, undisclosed to participants, there was also a second animacy rule, with two of the novel articles probabilistically predicting animate nouns and the other two predicting inanimate nouns. Using a gamified, speeded, performance-based task, participants were instructed that items and animals had "escaped" from their proper locations, and were asked to return them by clicking on the correct "home" for each item (e.g., a near zoo, a far shop). Critically, a small number of trials violated the hidden animacy rule (e.g., *gi lamp*, rather than *gi lion*). Following previous research by Batterink and colleagues (2014), our key prediction was that implicit learning of the novel animacy rule should produce delayed (as well as potentially less accurate) responses to violation trials relative to canonical trials. We considered reaction times as our primary dependent measure of rule sensitivity, given that is a finer-grained and generally more sensitive measure of processing than accuracy. We additionally included a small number of nonword trials, consisting of a novel article with a nonword (e.g., *gi badupi*) to assess generalization as well as explicit knowledge of the hidden rule during learning, reasoning that explicit insight into the hidden rule should result in a sudden jump in categorization accuracy for these trials. To assess offline consolidation over a 12-h delay period, participants performed the task over two sessions: an initial session in the evening, followed by a second session the next morning after a night of sleep. Finally, we assessed participants' explicit awareness of the hidden rule through a structured interview administered at the end of the second session. Our design therefore allowed us to compare children and adults' online sensitivity to a novel, hidden grammatical rule over multiple exposures, as well as their retention of this resulting knowledge over a delay.

2 | METHODS

2.1 | Participants

A total of 31 children (16 female; age range 8–10 years old; $M = 9.19$, SD = 0.99) and 30 adults (21 female; age range 18–35; $M = 24.65$, SD = 4.17) were initially recruited to participate. Five child participants were later excluded from analysis due to failure to perform the instructed experimental task (a performance metric i.e., independent of our main research contrasts; see Data Analysis section), leaving a final sample of 26 children. We set a goal of recruiting 30 participants per group prior to any confirmatory analyses being performed and in consideration of previous studies in this field (e.g., Batterink et al., 2014; Smalle et al., 2017). The inclusion criteria required that participants be native English speakers, have normal or corrected-to-normal vision and normal hearing, have no history of neurological or sleep disorders, and not be taking medication that may affect brain functioning. Informed consent was obtained from participants and parents, and assent was obtained from children. Participants were compensated for their time. The study was approved by the Research Ethics Board at the University of Western Ontario.

2.2 | Stimuli

As in Batterink and colleagues (2014), we used an artificial article system originally developed by Williams (2005). The article system consisted of four novel articles ("gi", "ro", "ul" and "ne"; see Figure 1a). Participants were instructed that these novel words functioned similarly to the word "the", with *gi* and *ro* indicating that the accompanying noun was near, and *ul* and *ne* indicating that the accompanying noun was far. Unbeknownst to participants, in addition to this explicit distance rule, there was a second "hidden" animacy rule: *gi* and *ul* typically preceded animals, while *ro* and *ne* typically preceded objects. Because Williams (2005) previously demonstrated that the specific assignment of animacy to each article did not affect learning, we kept animacy-article mappings consistent across participants.

Each trial contained a novel article paired with a unique noun (e.g., gi shirt = "the near shirt"). The nouns in this study consisted of 240 unique animal names and 240 unique object names. Selection of nouns was guided by the nouns that were used by Batterink and colleagues (2014) as well as age of acquisition ratings (Kuperman et al., 2012), with words that were associated with earlier age of acquisition selected for inclusion. Cartoon images of the objects and animals were sourced through Google Images, and edited to remove the background. An additional 80 nonwords were created using the ARC Nonword Database (Rastle et al., 2002), with settings selected to generate words that included only orthographically existing onsets, only orthographically existing bodies, only legal bigrams, and a range of 4–10 letters. All words (both articles and nouns) were recorded using a text-to-speech program (http://www.naturalreaders.com/) with speaker "Graham" at 0 speed. The audio was recorded and edited with Audacity software. All key test phrases during the main experimental task were presented

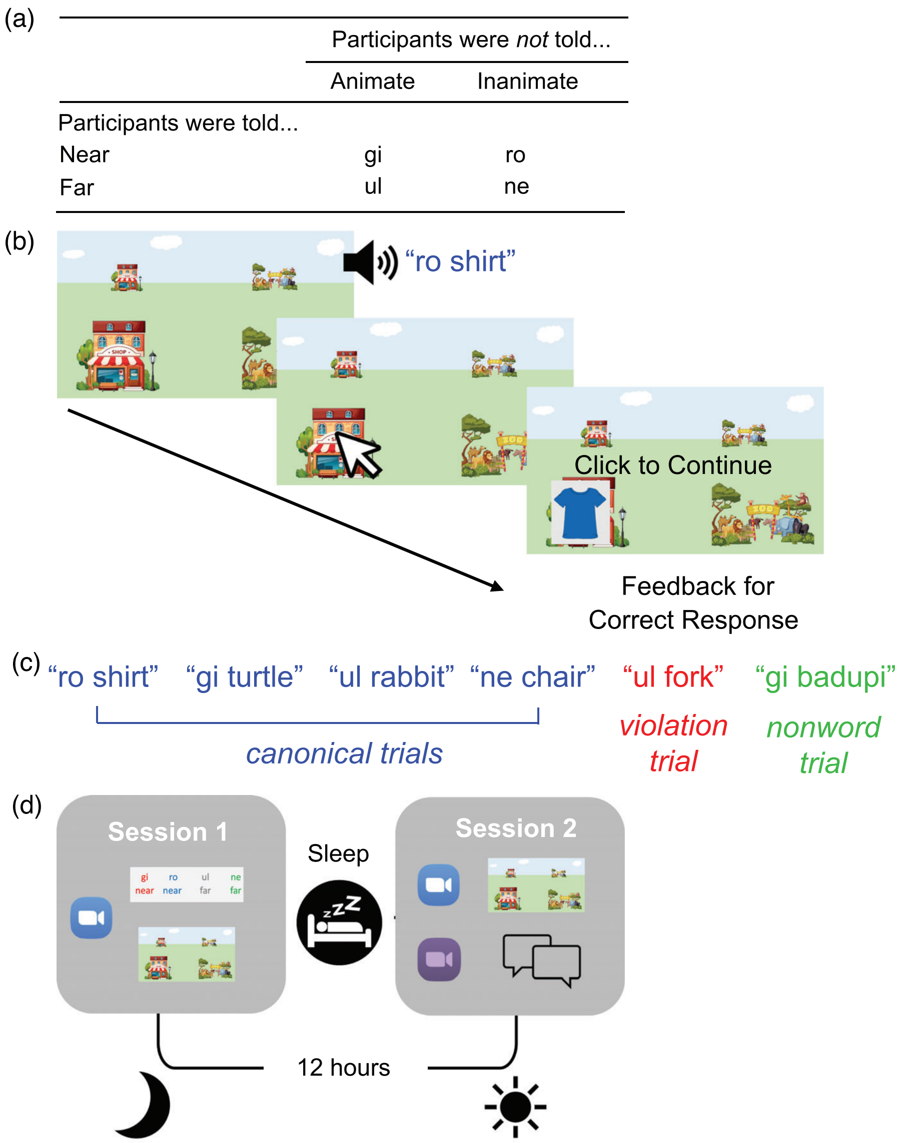

FIGURE 1 Summary of experimental task and overall procedure. (a) Summary of the miniature article system. (b) Sequence of events in a typical trial. (c) Representation of trial structure. Canonical trials (~71%) were interspersed pseudorandomly with less frequent violation trials (~14%) and nonword trials (~14%). (d) Each testing session included 280 trials. A 12-h delay period containing a period of nocturnal sleep separated the two sessions. Session 1 consisted of pretraining on the explicit distance rule, followed by the experimental task (conducted over Zoom for the children). Session 2 consisted of the experimental task, followed by an awareness assessment (conducted over Zoom for both children and adults).

auditorily, rather than through text as in our prior study (Batterink et al., 2014), to eliminate any effects of reading ability on task performance.

2.3 | Procedure

Before beginning the experiment, participants received a physical kit to take home, which contained electroencephalogram (EEG) and actigraphy equipment for monitoring sleep (sleep data not reported here). The experimental procedure consisted of an evening testing session (~1.5 h before participants' normal bedtime), an overnight in-home sleep session with portable EEG recording, and a morning session approximately 12 h after the first session. All experimental tasks were completed online, on participants' home computers. Participants first completed a questionnaire that included items relating to demographic information, language background, neurological history, vision and hearing, and current state/sleepiness via Qualtrics. For children, a Zoom call was then initiated with the experimenter prior to beginning the main experimental task. The experimenter remained on the Zoom call throughout the task to help guide and monitor the child. Adults were given the same set of instructions as children, but were not monitored over Zoom. The experimental task was created on PsychoPy software (Peirce et al., 2019) and administered through Pavlovia.org.

2.3.1 | Explicit pretraining tasks

In the first testing session only, participants first completed several pre-training tasks designed to encourage encoding of the explicit distance rule (i.e., that *gi* and *ro* meant "near"; *ul* and *ne* meant "far"). These tasks consisted of an initial memorization phase, in which participants were visually presented with the novel articles and their meanings and asked to memorize these pairings, followed by two pretraining tasks that required participants to match the articles to their correct meanings. Training proceeded until a pre-established criterion was met. For additional methodological details, please see Supporting Information.

2.3.2 | Hidden rule exposure task

Participants were then presented with a child-friendly cover story, in which objects and animals had "escaped" and the participant's help was needed to return each object and animal back to where it belonged. It was explained that the animals belonged in the zoo and the objects belonged in the shop. Participants were informed that, for each trial, they were to click on the near or far shop or zoo as fast as they could, and that if they failed to select the correct location for the item, they would hear the words repeated and have to try again. Therefore, both accuracy and speed were emphasized. The main experimental trials then began. As depicted in Figure 1b, each trial began with an image of two shops and two zoos, one of each which were "near" the participant (located towards the front of the screen) and the other which were "far" (located farther back on the screen). One of the four novel articles, followed by a noun, was then presented auditorily (e.g., "ro shirt"). The participant was asked to respond by clicking on the correct location for the item (i.e., the near or far shop or zoo), based on the distance of the novel article as well the animacy of the noun. If the participant clicked on the correct location, an image of the item appeared in the correct place, and the participant clicked to advance to the next trial. If the participant clicked on the wrong location, the phrase would be presented again. A correct response was required in order to move on to the next trial.

Both testing sessions included 280 trials each. Of these, 200 were canonical trials, in which the animacy of the noun corresponded to the hidden animacy rule (e.g., "ro shirt") and 40 were violation trials (10 per novel article), in which the animacy of the noun violated the hidden rule (e.g., "ro lion"). Finally, there were 40 nonword trials, in which the novel article was paired with a nonword (e.g., "ro badupi"; see Figure 1c). Before each nonword trial, a message was presented to inform participants that the next trial would contain a word that they had not heard before. Participants were instructed to sort these trials based on what they felt was best. Participants did not receive feedback on the nonword trials, and an image of a question mark appeared wherever they clicked.

Each session consisted of a different set of trials, such that each participant saw a given word or nonword just a single time. The trial order was pseudorandomized and ensured that violation trials, nonword trials and trials of the same article were distributed roughly evenly throughout the session. Specifically, 24 different preset pseudorandomized orders were created, and participants were assigned to a given order according to their participant ID. In addition, the objects and animals assigned to each session were counterbalanced, such that a given noun would be presented in session 1 for half the participants and in session 2 for the other half. The nouns were also counterbalanced to serve as violation trials across participants. Breaks were given every 40 trials, and length of the break was determined by the participant. The task lasted approximately 45 min.

2.3.3 | Structured awareness interview

At the end of the second session, a structured interview that probed participants' explicit knowledge of the hidden animacy rule was conducted over Zoom. The experimenter asked a series of questions and immediately transcribed participants' responses. The questions became more specific as the interview went on. Participants were first asked more open-ended questions to probe whether they had any knowledge of the hidden animacy rule. If they failed to describe animacy as a relevant feature, they were then directly asked whether they thought the novel articles had anything to do with the item being an object or animal, and if so, to describe. Participant responses were later coded as aware or unaware. Based on low overall levels of participant awareness, we used a liberal awareness criterion (and thus a conservative "unaware" criterion), coding participants as aware if they correctly indicated the relevance of animacy for at least one of the four novel articles at any point in the interview (e.g., "ul was for animals" or "ro was for objects").

2.4 | Data analyses

All analyses were conducted with R software (R Core Team, 2020), mixed effects models were conducted using the *lme4* package (Bates et al., 2015; Lenth et al., 2022), and interactions were interpreted using the *emmeans* package (Lenth et al., 2022). Participants who performed below chance overall on the canonical trials (i.e., <25% accuracy for first responses) were excluded due to failure to perform the explicitly instructed task (n = 5 children). This resulted in a final sample of 26 children and 30 adults for all analyses. At the trial level, we excluded outliers by applying the following steps. First, any trial with a response time (RT) greater than 60 s was excluded (eight trials in total). While rare, these extremely long responses reflected occasional instances in which the participant was interrupted during the task or took an otherwise unscheduled break; inclusion of these trials would result in inflated standard deviations (SDs) for several participants. Next, for each participant, all trials with a RT less than or greater than 3 SDs from the participant's mean RT were discarded (1.4% of trials in children; 1.7% of trials in adults). For all analyses, we fitted maximal random effects warranted by the data (Barr et al., 2013).

2.4.1 | Emergence and time-course of sensitivity to the implicit animacy rule

Given that RTs provide a more fine-grained measure of accuracy, and following previous research by Batterink and colleagues (2014), we considered RT delays to violation trials relative to canonical trials to be our primary index of sensitivity to the hidden animacy rule. Therefore, our first analysis examined whether RTs to correct trials may differ between canonical and violation trials. As a secondary measure, we also tested whether violation trials were responded to less accurately than canonical trials.

For the RT analysis, because children's RTs were slower than adults, RTs for correct trials (i.e., only trials for which participants gave an initial correct responses) were within-subject *z*-scored to allow for more direct RT comparisons between groups. A linear mixed effects model was conducted on *z*-normalized RTs, with condition (canonical vs. violation), trial number (1–560; scaled to support model convergence), session (1,2), age group (child, adult), and their factorial interactions as fixed factors, and item intercept as a random effect, which represents the maximal random effects structured warranted by the data. Inclusion of other random effects and slopes was explored but resulted in the model obtaining a singular fit. Treatment coding was used for condition (with canonical as the reference category) and session (with session 1 as the reference; thus, effect of trial was estimated within the first session, allowing for characterizing earlier phases of learning at a finer-grained resolution). Sum coding was used for age group, such that effects of other variables are reported across all participants (cf. Brehm & Alday, 2022). We expected RT to generally decrease over time, reflecting improved fluency with the task with additional exposure and practice. More critically, we expected slower RTs for violation trials compared to canonical trials as time went on, reflecting learning of the hidden animacy rule.

For accuracy, a generalized mixed effects logistic regression was conducted on accuracy for each trial (1 = correct, 0 = incorrect) with trial number (1–560; scaled to support model convergence), session (1,2), age group (child, adult), and their factorial interactions as fixed factors, and participant and item intercept as random effects, with the optimizer set to bobyqa. As with RTs, treatment coding was used for condition (with canonical as the reference) and session (with session 1 as the reference), and sum coding was used for age group. Similarly to RT, we expected that accuracy should increase overall across the two sessions, with lower accuracy for violation trials relative to canonical trials.

We additionally used a single, combined measure of speed and accuracy called the Balanced Integration Score (BIS), which controls for potential speed accuracy trade-offs (Liesefeld et al., 2015), to thoroughly characterize the emergence of hidden rule sensitivity across learning blocks within each age group. This analysis allowed us to quantify approximately when during learning each group showed reliable overall sensitivity to the rule, considering both speed and accuracy in combination. This measure is calculated by *z*-scoring reaction times and the percentage of correct trials, then subtracting the standardized RTs from the standardized percent correct (Liesefeld & Janczyk, 2019). Because the BIS is a composite measure that cannot be modeled at the trial level, we separated each session into two blocks (each containing 120 trials), resulting in four blocks total across both sessions (i.e., blocks 1_A, 1_B, 2_A, and 2_B; see Batterink et al., 2014, for a similar four block time-course approach). Reaction times and percentage of correct trials were *z*-scored separately for each age group and the BIS was calculated by condition, participant, and block. Given adults' overall more accurate and faster responses, we *z*-scored within each group rather than across groups, which remove the main effect of group on performance (cf. Liesefeld & Janczyk, 2019) and allows for visualizing the time course of learning for each group on comparable scales.

To establish that BIS was a sensitive index of violation processing, we first ran a basic model on BIS scores, including data from all four blocks, with condition (canonical vs. violation), age group and their interaction as fixed effects and participant as a random effect. Sum coding was used for age group and treatment coding was used for condition (with canonical as the reference). We expected to see overall higher BIS values for canonical trials than violation trials, indicating relatively facilitated performance. Applying this same model with all the same factors just described, we then conducted targeted analyses within each session (session 1 model: data from blocks 1_A and 1_B; session 2 model; data from blocks 2_A and 2_B) to further test the hypothesis that children would show earlier-emerging sensitivity to the rule within the first session when both accuracy and RT are considered. Finally, to characterize the overall time course of learning and to pinpoint the earliest emergence of rule sensitivity within each group, we conducted a series of paired samples t-tests between the canonical and violation conditions, within each block (1_A, 1_B, 2_A, 2_B) and age group. We predicted that a difference between canonical and violation trials would emerge earlier in children than in adults.

2.4.2 | Generalization of hidden rule (nonword trials)

To isolate generalization of the hidden animacy rule specifically, we excluded all nonword trials in which the explicit distance rule was incorrectly applied (children: 14.5%; adults: 6.6%). A mixed effects logistic regression was then conducted on the remaining trials with animacy accuracy (determined by the article) as the dependent measure (1 = correct, 0 = incorrect), age group as a between-subjects factor, session (1,2) as a within-subject factor, and participant and item as a random effect. The trial number was not included due to model fit issues. Sum coding was used for age group and session, such that effects of age refer to main effects across both sessions.

2.4.3 | Overnight retention of rule knowledge (BIS values)

To characterize the impact of the delay period on knowledge of the hidden animacy rule, we ran an additional model to test for significant

TABLE 1 Results of normalized RT analysis (correct trials only).

Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.14	0.09–0.18	**<0.001**
Condition[T.Violation]	0.07	−0.04–0.18	0.197
Trial	−0.34	−0.48–−0.19	**<0.001**
Session[T.2]	−0.50	−0.62–−0.39	**<0.001**
Age group[S.Adult]	0.10	0.06–0.14	**<0.001**
Condition[T.Violation] ×Trial	0.06	−0.31–0.43	0.733
Condition[T.Violation] ×Session[T.2]	−0.25	−0.55–0.05	0.097
Trial ×Session[T.2]	0.73	0.52–0.94	**<0.001**
Condition[T.Violation] × AgeGroup[S.Adult]	−0.11	−0.21–−0.01	**0.040**
Trial × AgeGroup [S.Adult]	−0.29	−0.44–−0.15	**<0.001**
Session[T.2] × AgeGroup[S.Adult]	0.24	0.12–0.36	**<0.001**
Condition[T.Violation] ×Trial ×Session[T.2]	0.26	−0.25–0.78	0.317
Condition[T.Violation] ×Trial × AgeGroup [S.Adult]	0.46	0.10–0.83	**0.013**
Condition[T.Violation] ×Session[T.2] × AgeGroup [S.Adult]	0.15	−0.14–0.45	0.306
Trial ×Session[T.2] × AgeGroup[S.Adult]	−0.20	−0.41–0.00	0.052
Condition[T.Violation] ×Trial × Session[T.2] × AgeGroup [S.Adult]	−0.56	−1.07–−0.04	**0.033**

Note: Labels within square brackets indicate treatment or sum coding ["T" or "S"] and the level associated with the positive value. Bolded parameters indicate significance. Model syntax: lmer(normRT ~ Condition × Trial × Session × AgeGroup + (1|item)).

effects of the preceding delay period over and above general effects of trial number. If the opportunity for consolidation benefits knowledge of the implicit animacy rule, we would expect to observe a sudden increase in sensitivity to the animacy rule immediately following the delay period, dissociable from effects of additional exposure to the task. Thus, for this analysis, comparison of trials immediately following the delay period relative to other trials were of particular interest. We created a "post-delay" categorical variable, coding the block immediately following the delay period as 1 (block 2_A, representing trials that occurred immediately post-delay), with the other three blocks coded as 0. We then conducted a linear mixed effect model on BIS values that included interactions of block (1–4), age group and condition as fixed factors, in addition to interactions between immediate-post-delay category (0,1), age group and condition, with participant intercept again modeled as a random effect. Sum coding was again used for age group and treatment coding was used for condition (with canonical as the reference). A significant interaction between the immediate-post-delay variable and condition would indicate that participants showed differential sensitivity to the animacy rule immediately after the delay period relative to the other three blocks, which is over and above the general effects of continued exposure to the task.

Using a similar logic, we also tested whether there was a significant impact of the delay period on knowledge of the explicitly-instructed distance rule by examining BIS values for canonical trials only. We conducted a linear mixed effect model that included block (1–4), age group and their interaction, as well as the interaction between immediate-post-delay and age group. Age group was again sum coded, and participant intercept was again modeled as a random effect. A significant effect of immediate-post-delay would indicate that participants' explicit categorization performance immediately after the delay period showed a divergence in performance from what would be expected based on effects of additional exposure alone.

3 | RESULTS

3.1 | Emergence and time-course of sensitivity to the hidden animacy rule

3.1.1 | RT analysis

RTs served as our primary dependent measure of rule learning. Our main model indicated that, across both age groups, RTs to canonical trials became faster over time (Trial [within the first session]: $p < 0.001$; Session: $p < 0.001$; see Table 1). Within the first session, the effect of condition was not significant ($p = 0.20$), which may be attributed to the large amount of variance accounted for by trial and session in this large model. A simpler model with only condition and age group as fixed effects and item as a random effect revealed a main effect of condition across the experiment (estimate = 0.073, CI = 0.04–0.11, $p < 0.001$, where violation trials had slower RTs than canonical trials. This model revealed no main effect of age group (estimate = 0.00, CI = −0.01–0.02, $p = 0.94$), nor interaction between age group and condition (estimate = −0.01, CI = −0.05–0.03, $p = 0.65$). These results indicate that, when considering responses across the two sessions, both age groups showed overall sensitivity to the hidden rule.

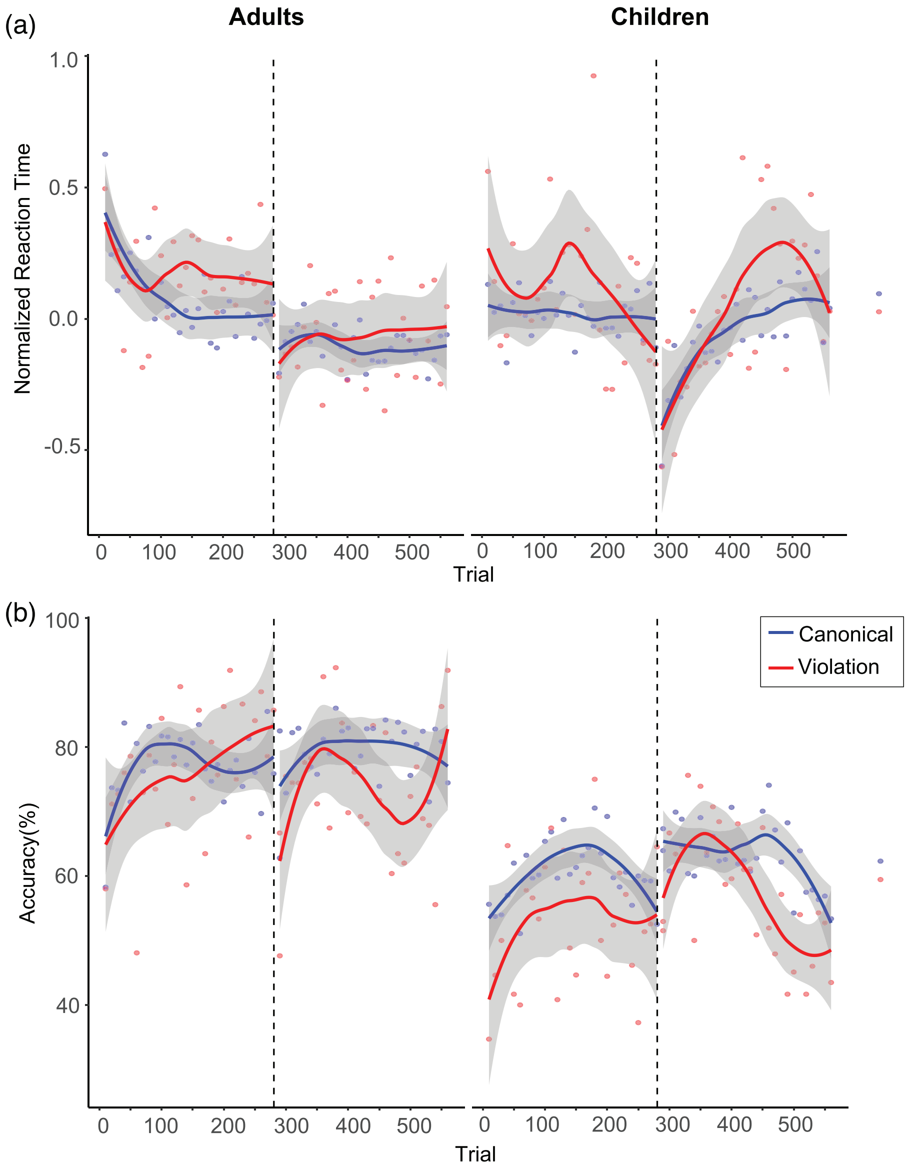

FIGURE 2 Mean (a) normalized reaction time and (b) accuracy percentage, for canonical and violation trials, averaged across 10-trial bins. Each dot represents the group mean within a particular bin. The curved lines represent the smoothed trend for each group, session, and condition, with the shaded error bars representing the 95% confidence interval. The vertical dotted line represents the overnight break during which sleep occurred. This division into bins is used for visualization purposes only, as our statistical analyses modeled responses at the individual trial level.

Returning to the main model, the two age groups showed differences in the progression of the violation effect within the first session, as indicated by a significant three-way interaction of condition, trial and age group ($p = 0.013$). To follow up on this interaction, we examined the separate linear trends of trial number by condition within each age group. In adults, there was a significant difference in slopes between canonical and violation trials across time (contrast estimate [canonical—violation] = −0.38, SE = 0.17; z ratio = −2.25, $p = 0.025$), reflecting relatively greater facilitation over time for canonical trials. In contrast, children showed *no* significant difference in slopes between canonical and violation trials over time (contrast estimate [canonical—violation] = −0.013, SE = 0.20; z ratio = −0.064, $p = 0.95$). As illustrated in Figure 2a, these results indicate that adults became increasingly sensitive to violations of the animacy rule as the first session progressed, whereas in children the violation effect was present from very early on in learning, remaining relatively stable thereafter.

3.1.2 | Accuracy analysis

Accuracy served as an additional, secondary measure of rule sensitivity. Overall, as expected, participants performed more accurately for canonical than violation trials (canonical M = 70.4%, violation M = 64.0%; $p < 0.001$; Table 2), providing evidence of learning the hidden rule. Unsurprisingly, there was also a significant effect of age group ($p < 0.001$), indicating that children showed significantly poorer

TABLE 2 Results of accuracy analysis.

Predictors	*Odds Ratios*	*CI*	*p*
(Intercept)	2.90	2.31–3.64	**<0.001**
Condition[T.Violation]	0.67	0.55–0.81	**<0.001**
Trial	1.18	0.88–1.59	0.258
AgeGroup[S.Adult]	1.60	1.28–1.99	**<0.001**
Session[T.2]	1.37	1.07–1.76	**0.013**
Condition[T.Violation] × Trial	2.02	1.00–4.08	**0.050**
Condition[T.Violation] × AgeGroup[S.Adult]	1.00	0.83–1.21	0.983
Trial × AgeGroup[S.Adult]	1.03	0.78–1.38	0.815
Condition[T.Violation] × Session[T.2]	2.18	1.22–3.92	**0.009**
Trial × Session[T.2]	0.68	0.44–1.04	0.073
AgeGroup[S.Adult] × Session[T.2]	0.68	0.53–0.86	**0.002**
Condition[T.Violation] × Trial × AgeGroup[S.Adult]	1.81	0.91–3.60	0.092
Condition[T.Violation] × Trial × Session[T.2]	0.20	0.07–0.53	**0.001**
Condition[T.Violation] × AgeGroup[S.Adult] × Session[T.2]	0.82	0.46–1.45	0.491
Trial × AgeGroup[S.Adult] × Session[T.2]	1.63	1.08–2.47	**0.021**
Condition[T.Violation] × Trial × AgeGroup[S.Adult] × Session[T.2]	0.69	0.26–1.84	0.459

Note: Labels within square brackets indicate treatment or sum coding ["T" or "S"] and the level associated with the positive value. Bolded parameters indicate significance. Model syntax: Glmer(Accuracy ~ Condition × Trial × AgeGroup × Session +(1|participant) + (1|noun), family = "binomial", glmerControl(optimizer = "bobyqa", optCtrl = list(maxfun = 100000))).

accuracy (M = 60.4%) than adults (M = 77.5%). In addition, accuracy for canonical trials improved from the first session to the second session (p = 0.013). Regarding our main hypothesis about differences in the progression of the violation effect, the condition × trial × age group interaction did not reach significance (p = 0.092). However, we note that the accuracy data numerically followed a similar pattern as the RT data, with children showing a larger difference between canonical and violation trials than adults during early learning stages (Figure 2b).

3.1.3 | BIS analysis

The RT results—our primary measure of hidden rule learning—indicate that children became sensitive to the animacy rule at an earlier stage than adults. To further characterize the emergence of hidden rule sensitivity within each age group, we conducted an additional analysis using the BIS measure, computed within each of our four previously-defined blocks. In a full model with condition, block, and age group as fixed effects and participant as a random effect, no predictors were significant (Table 3), likely due to the large amount of variance account for by block. A reduced model with condition and age group as fixed effects and participant as a random effect revealed a highly significant effect of condition across the four blocks (estimate = 0.35, SE = 0.075, $t(390)$ = 4.68, p < 0.001), such that canonical trials were associated with higher BIS scores (indicative of facilitated performance; Table 4). This result confirms that the BIS measure captures participants' overall sensitivity to the hidden rule. Children also showed a marginally greater BIS overall condition effect than adults (Condition × Age Group: p = 0.087)

TABLE 3 Results of BIS analysis (full model).

Predictors	*Estimates*	*CI*	*p*
(Intercept)	−0.34	−0.77–0.08	0.116
Condition[T.Canonical]	0.26	−0.10–0.62	0.152
Block	0.07	−0.03–0.16	0.158
AgeGroup[S.Adult]	0.00	−0.43–0.43	1.000
Condition[T.Canonical] × Block	0.04	−0.10–0.17	0.594
Condition[T.Canonical] × AgeGroup[S.Adult]	−0.25	−0.61–0.11	0.177
Block × AgeGroup[S.Adult]	0.03	−0.07–0.12	0.585
Condition[T.Canonical] × Block × AgeGroup[S.Adult]	0.05	−0.08–0.18	0.479

Note: Labels within square brackets indicate treatment or sum coding ["T" or "S"] and the level associated with the positive value. Bolded parameters indicate significance. Model syntax: Lmer(bis ~ Condition × Block × AgeGroup + (1 | participant).
Abbreviation: BIS, balanced integration score.

Given the RT finding that children became sensitive to the hidden rule more quickly than adults within the first session, we conducted targeted analyses to further test this effect while ruling out speed-accuracy trade-offs, taking into account both RT and accuracy simultaneously, as provided by the BIS measure. Within the first session (blocks 1_A and 1_B), children showed a significantly greater overall

TABLE 4 Results of BIS analysis (reduced model).

Predictors	*Estimates*	*CI*	*p*
(Intercept)	−0.18	−0.54–0.18	0.337
Condition[T.Canonical]	0.35	0.20–0.50	**<0.001**
AgeGroup[S.Adult]	0.06	−0.30–0.42	0.725
Condition[T.Canonical] × AgeGroup[S.Adult]	−0.13	−0.28–0.02	0.087

Note: Labels within square brackets indicate treatment or sum coding ["T" or "S"] and the level associated with the positive value. Bolded parameters indicate significance. Model syntax: lmer(bis ~ Condition × AgeGroup + (1 | participant).
Abbreviation: BIS, balanced integration score.

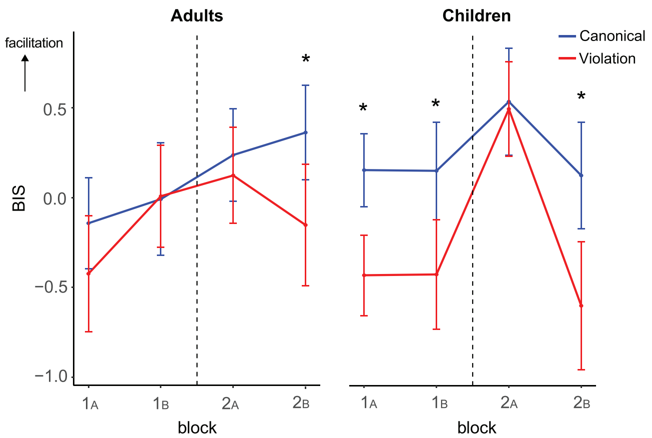

FIGURE 3 The BIS, an integrated measure of speed and accuracy, for canonical and violation trials, plotted over the four blocks for each age group. More positive values indicate relative facilitation (higher accuracy and faster response times). Children showed a significant violation cost from very early on, whereas the violation effect in adults did not reach significance until the 4th block (asterisk indicates $p < 0.05$, paired sample t-test between conditions within each block). Error bars represent standard error. BIS, balanced integration score.

violation effect than adults (condition × age group: estimate = −0.22, SE = 0.088; $t(166) = 2.56$, $p = 0.011$). Follow up analyses within the first two blocks indicated that only children showed a significant condition effect during the first session; the condition effect in adults was not significant (children: estimate = 0.58, SE = 0.13; $t(77) = 4.32$, $p < 0.001$; adults: estimate = 0.13, SE = 0.11; $t(89) = 1.16$, $p = 0.25$; see Figure 3). In contrast, by the second session (blocks 2_A and 2_B), both children and adults showed a significant, and similar, violation effect (condition: estimate = 0.35, SE = 0.11; $t(166) = 3.32$, $p = 0.001$; condition × age group: estimate = −0.034, SE = 0.11; $t(166) = -0.33$, $p = 0.74$). The BIS results thus provide additional evidence that children became reliably attuned to the hidden rule at an earlier point in time than adults.

Finally, we used the BIS measure within each of the four blocks to more precisely quantify *when* in the learning process each age group showed reliable sensitivity to the hidden rule, providing an overall picture of the time course of learning that incorporates both RT and accuracy. As expected, based on the prior RT analyses, reliable sensitivity to the rule emerged early on in children. Children showed a

TABLE 5 Results of generalization analysis (nonword trials only).

Predictors	*Odds ratios*	*CI*	*P*
(Intercept)	0.92	0.78–1.08	0.291
AgeGroup[S.Adult]	0.86	0.79–0.94	**0.001**
Session[S.1]	0.93	0.87–0.99	**0.034**
AgeGroup[S.Adult] × Session[S.1]	0.99	0.93–1.06	0.809

Note: Labels within square brackets indicate treatment or sum coding ["T" or "S"] and the level associated with the positive value. Bolded parameters indicate significance. Model syntax: Glmer(correct ~ AgeGroup × Session + (1 | participant) +(1 | item), family = "binomial").

significant difference in BIS between canonical and violation trials in block 1_A($t(25) = 3.44$, $p = 0.002$), block 1_B ($t(25) = 4.61$, $p = 0.001$), and block 2_B ($t(25) = 3.70$, $p = 0.001$; see Figure 3). Interestingly, the violation effect was not significant in block 2_A ($t(25) = 0.18$, $p = 0.86$), suggesting a transient reduction in implicit sensitivity to the animacy rule immediately following the 12-h consolidation period. In contrast, adults showed a significant condition difference only in block 2_B (i.e., past the halfway point of the second session; $t(29) = 2.98$, $p = 0.006$; all p values for previous blocks ≥ 0.1). This analysis confirms that children showed sensitivity to the violation rule at an earlier point in learning.

3.2 | Generalization of hidden rule (nonword trials)

While generalization performance on nonword trials was generally low, children showed overall significantly higher generalization performance than adults across both sessions ($p = 0.001$; Table 5). In addition, performance improved across both age groups from session 1 to session 2 ($p = 0.034$). There was no significant interaction between age group and session ($p = 0.81$).

One-tailed t-tests were then conducted to test whether participants' animacy rule generalization performance was above chance (50%) within each of the four blocks. Children performed significantly better than chance during block 2_B (the second half of the second session), $t(442) = 1.76$, $p_{one\ tailed} = 0.039$, but did not achieve above-chance accuracy on any of the other blocks ($p_{one\text{-}tailed} > 0.15$). Adults did not perform above chance in any of the four blocks (all p values > 0.9; Figure 4).

In fact, adults (as well as children during block one) performed below chance on the first three blocks, which was unexpected. We explored this result in a supplementary analysis and found evidence that this below-chance performance was related to an unintended in animacy bias shown by participants for the articles *gi* and *ul* (which actually correspond to animate items). That is, despite the fact that the novel words were randomly assigned to the articles and counterbalanced, for *gi* and *ul*, participants selected the shop more frequently than would be expected by chance ($p < 0.001$). We speculate that some particular acoustic feature(s) of these words may have led participants to associate them with the concept of nonliving (e.g., D'Anselmo et al., 2019).

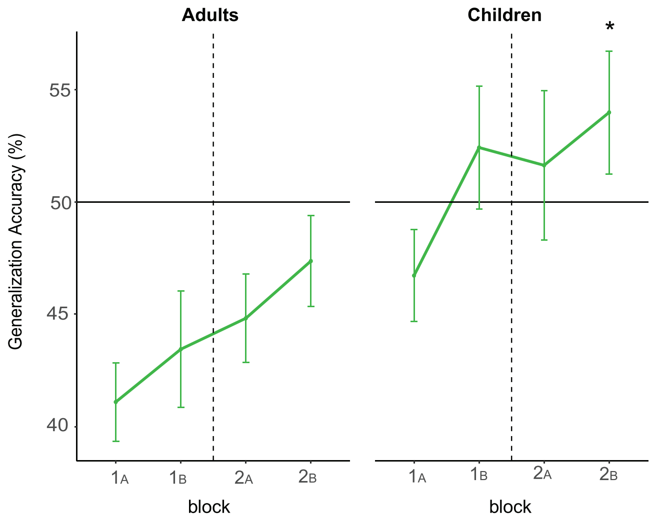

FIGURE 4 Animacy accuracy for nonword trials, including only trials in which participants made a correct explicit distance judgement. Chance is 50%. Only children in the final block achieved above-chance accuracy. Error bars represent standard error.

Please see Supplementary Information for more details of this analysis and further discussion of this issue.

3.3 | Rule awareness and its effect on task performance

A relatively low number of participants were able to report any verbalizable knowledge about the hidden animacy rule (38% children; 23% adults). Although a numerically greater proportion of children demonstrated awareness of the hidden rule than adults, this difference was not significant, X^2 (1, N = 56) = 1.51, p = 0.219.

We conducted several analyses designed to test the impact of rule awareness on task performance. Our main findings, showing a cost in violation processing and earlier emergence of hidden rule sensitivity in children compared to adults, were maintained even after excluding any participant who gained conscious awareness of the rule (see Supporting Information for additional information on the role of awareness on task performance).

3.4 | Retention across the delay period

Across both age groups, as indexed by BIS values, participants showed a *reduced* sensitivity to the hidden animacy rule in the trials immediately following the delay period (block 2_A) relative to the other three blocks, over and above effects of continued exposure to the task (Post-Delay × Condition: p = 0.015; Table 6). This delay-related reduction in the violation effect did not significantly differ between children and adults (Post-delay × Condition × Age group: p = 0.24). Follow-up contrasts showed that the violation effect in block 2_A was not significant (estimate = 0.038, SE = 0.15, t(382) = 0.26, p = 0.79), whereas

TABLE 6 Effect of delay on hidden animacy rule (BIS values).

Predictors	Estimates	CI	p
(Intercept)	−0.34	−0.77–0.08	0.111
Block Post Delay	0.64	0.41–0.88	**<0.001**
AgeGroup[S.Adult]	0.00	−0.42–0.42	1.000
Condition[T.Canonical]	0.26	−0.08–0.61	0.133
Block	0.00	−0.09–0.09	0.957
Block Post Delay × AgeGroup[S.Adult]	−0.38	−0.61–−0.14	**0.002**
Block Post Delay × Condition[T.Canonical]	−0.42	−0.75–−0.08	**0.015**
AgeGroup[S.Adult] × Condition[T.Canonical]	−0.25	−0.59–0.10	0.158
AgeGroup[S.Adult] × Block	0.06	−0.03–0.16	0.176
Condition[T.Canonical] × Block	0.08	−0.05–0.21	0.241
Block Post Delay × AgeGroup[S.Adult] × Condition[T.Canonical]	0.20	−0.13–0.54	0.235
AgeGroup[S.Adult] × Condition[T.Canonical] × Block	0.03	−0.10–0.16	0.682

Note: Labels within square brackets indicate treatment or sum coding ["T" or "S"] and the level associated with the positive value. Bolded parameters indicate significance. Model syntax: Lmer(bis ∼ Block_post_delay × AgeGroup*Condition +Block × AgeGroup × Condition + (1 | participant).
Abbreviation: BIS, balanced integration score.

it was significant across the other three blocks (estimate = 0.45, SE = 0.083, t(382) = 5.49, p < 0.001). Altogether, these results suggest that participants in both age groups showed a transient reduction or destabilization in their representation of the hidden animacy rule after the 12-h delay period.

Next, we examined consolidation of the explicit distance rule by analyzing BIS values for canonical trials only. Across both age groups, performance was facilitated for the block immediately following the delay period relative to the other three blocks (Post-Delay: p = 0.016; Table 7), which occurred over and above the general facilitation by block (Block: p = 0.027). In addition, this delay-related boost in performance differed marginally between children and adults (p = 0.063). Follow-up contrasts showed that children showed a significant delay-related boost in explicit categorization performance, over and above effects of additional exposure (estimate = 0.40, SE = 0.14, t(164) = 2.95, p = 0.004; Figure 3). In contrast, adults did not show a significant boost in performance for trials that occurred immediately post-delay (estimate = 0.053, SE = 0.13, t(164) = −0.42, p = 0.68).

4 | DISCUSSION

Our results demonstrate that children possess advantages in real-time language learning, supporting our primary hypothesis. While both children and adults gained sensitivity to the hidden linguistic rule,

TABLE 7 Effect of delay on instructed distance rule (BIS values, canonical trials only).

Predictors	*Estimates*	*CI*	*P*
(Intercept)	−0.08	−0.48–0.31	0.688
Block Post Delay	0.23	0.04–0.41	**0.016**
AgeGroup[S.Adult]	−0.25	−0.64–0.15	0.218
Block	0.08	0.01–0.15	**0.027**
Block Post Delay × AgeGroup[S.Adult]	−0.17	−0.36–0.01	0.063
AgeGroup[S.Adult] × Block	0.09	0.02–0.16	**0.013**

Note: Labels within square brackets indicate treatment or sum coding ["T" or "S"] and the level associated with the positive value. Bolded parameters indicate significance. Model syntax: Lmer(bis ~ Block_post_delay × AgeGroup + Block × AgeGroup + (1 | participant).
Abbreviation: BIS, balanced integration score.

as demonstrated by slower RTs and decreased accuracy to violation trials compared to canonical trials, we found that sensitivity to the hidden rule emerged more quickly in children than adults. Children showed a processing cost for violation trials from very early on in learning, whereas adults did not show a violation cost initially but became increasingly more sensitive to the hidden rule as the task progressed. On our integrated measure of speed and accuracy (the BIS), the violation cost in children was significant from the first block, but reached significance in adults only in the final block. This overall violation cost was also significantly greater in children than adults within the first session, and marginally greater across both sessions. Children also outperformed adults on generalization of the hidden animacy rule, eventually classifying nonword trials along their hidden animacy dimension at above-chance levels. In contrast, adults as a group failed to generalize the hidden animacy rule to nonwords at any point during the task. Overall, these findings suggest that children's superior language attainment is partially driven by their ability to more rapidly extract hidden linguistic structures from input during real-time language exposure.

4.1 | Children demonstrate linguistic rule learning advantages over adults

Our results provide strong support for the idea that children have a true advantage for learning linguistic regularities—not only in terms of their ultimate attainment levels, but also in terms of their *rate* of learning. While the idea that children are faster language learners than adults represents a pervasive popular belief, there is surprisingly little empirical evidence to directly support this view. As mentioned previously in the Introduction, most studies comparing adults and children's learning under equivalent conditions show that adults—not children—learn language more quickly during initial learning stages (Asher & Price, 1967; Ferman & Karni, 2010; Hudson Kam & Newport, 2009; Krashen et al., 1979; Lichtman, 2016; Long, 1990; Snow & Hoefnagel-Höhle, 1978); the study by Smalle and colleagues (2017) represents a recent exception). Nonetheless, because children eventually surpass adults in the acquisition of grammar and phonology, both classic and current theories have suggested that children rely on different and possibly more efficient learning mechanisms that ultimately enable them to better acquire implicit or procedural aspects of language (e.g., DeKeyser, 2000; Kareev, 1995; Lenneberg et al., 1967; Newport, 1990; Paradis, 2009; Thiessen et al., 2016; Ullman, 2001, 2004, 2005). Our results provide novel evidence in support of these views, showing that there are developmental differences in real-time learning ability underlying the acquisition of abstract, hidden linguistic structures.

The current paradigm incorporated an explicitly instructed rule (i.e., the distance rule) alongside an undisclosed rule that was not directly relevant to the assigned experimental task (i.e., the animacy rule). Adults showed a clear advantage in classifying phrases according to the instructed rule (as demonstrated by their overall better accuracy as well as faster reaction times), while children become sensitive to the hidden, uninstructed rule more quickly. This age-related dissociation closely resembles a number of related findings showing that children process and learn task-irrelevant information better than adults (e.g., Blanco & Sloutsky, 2019; Decker et al., 2015; Jung et al., 2023; Plebanek & Sloutsky, 2017). In one study, young children (ages 4–5) and adults were asked to perform a visual search task in which stimuli had a task-relevant dimension and task-irrelevant dimensions, and then tested on their memory for these stimuli. Compared to adults, children showed better recognition of stimuli with irrelevant features (Plebanek & Sloutsky, 2017). In a related study, participants engaged in a probabilistic learning task, in which they were falsely instructed that one stimulus had a high likelihood of being rewarded, when in actuality it did not. Adults showed a stronger bias towards selecting the stimulus with the false instruction, while adolescents (13–17 years) and children (6–12 years) more quickly abandoned this incorrect information as they learned the true probabilities through experience (Decker et al., 2015). Recent fMRI evidence underscores these results, showing that children (7–9 years) represent both task-relevant and task-irrelevant information in visual cortex to a similar extent, unlike adults who prioritize task relevant information much more strongly (Jung et al., 2023). Taken together, these results suggest that adults may be at times "hyper-focused" on the task at the hand, at the cost of processing information that is external to their assigned task. In contrast, children's more diffuse focus of attention may lead them to process irrelevant information more deeply, which may in turn result in greater incidental learning of extraneous information. The current results can clearly be understood within such a framework. Adults in the current study may have focused narrowly on the assigned task (i.e., "sort items according to the distance rule"), whereas children may distribute their attention more broadly, thereby processing seemingly irrelevant information that enabled them to better extract the hidden, secondary rule.

The current findings also closely parallel those from a recent study that compared children and adults' learning of phonotactic constraints (Smalle et al., 2017). In that study, 9–10-year-old children showed reliable evidence of learning second order phonotactic constraints within the first day of training, after exposure to only 24 novel word-form sequences. In contrast—and as in the current study—adults only showed evidence of learning by the second session, after a delay period containing sleep. Additional findings from adult studies confirm the idea that sleep may promote, or even be necessary, for phonotactic learning to occur past childhood. Gaskell et al. (2014) found that adults who slept, but not those who stayed awake, showed evidence of learning phonotactic constraints. Another study in adults tested whether a period of consolidation benefits phonotactic constraint learning over and above more exposure to the regularities (Warker, 2013). The authors found that a consolidation period resulted in a greater learning benefit than a longer initial training session. Our results extend these findings from the phonotactic constraint literature to a novel linguistic paradigm, supporting the idea that children can learn linguistic rules after only a brief period of exposure, whereas adults may require additional exposure and/or a period of consolidation to stabilize this implicit knowledge.

An additional parallel can be drawn to changes in motor learning and perceptual learning that occur across development. Evidence from motor learning paradigms suggests that after initial learning, children's procedural memories stabilize very rapidly and can then be immediately expressed as improved performance (Adi-Japha et al., 2014; Ashtamker & Karni, 2013). In contrast, adults' procedural memories remain susceptible to inference for several hours after learning and stabilize over a longer time period (Adi-Japha et al., 2014; Dorfberger et al., 2007), a process that may depend on sleep (Korman et al., 2007). Similarly, it was also recently shown that children's visual perceptual learning stabilizes very rapidly after training and is highly resilient to retrograde interference, whereas adults' perceptual learning is much more fragile, remaining highly susceptible to interference for at least 1 h after training. This rapid stabilization is children is supported by a rapid boost of GABA in visual cortex during training, an effect not seen in adults (Frank et al., 2022). In the current paradigm, the intermittent violation trials may be considered a form of interference, and may have more strongly disrupted adults' ability to extract the hidden animacy rule during real-time learning, relative to children. One intriguing proposed idea is that the mechanisms underlying rapid procedural memory consolidation in children can be recruited by children in the awake state, but are engaged by adults only during sleep (Adi-Japha et al., 2014; Wilhelm, Prehn-Kristensen, Born, 2012). Such an idea fits well with the current findings, where reliable evidence of rule sensitivity emerged early on during learning in children, but not until after a period containing sleep in adults (see BIS analysis). However, we should note that our design does not allow us to directly disentangle the effects of additional exposure to the rule from effects of consolidation. Both additional rule exposure and the opportunity for consolidation may have contributed to the eventual emergence of implicit rule sensitivity in adults during the second session.

In addition to showing overall greater sensitivity to the hidden rule, children also showed better rule generalization performance than adults, more accurately indicating the correct animacy when a novel article was presented with a meaningless nonword (e.g., *ro badupi*). While generalization performance in both groups was generally poor, children did eventually achieve above chance generalization performance in the final block, whereas adults failed to reach above-chance generalization performance at any point. This result provides converging evidence of children's advantage for learning hidden linguistic structures.

Along with assessing generalization ability, nonword trials were also initially intended to provide an online index of rule awareness; we reasoned that participants who gained explicit insight into the hidden rule during learning would show a sudden jump in accuracy on these trials, performing at or near ceiling. Consistent with this reasoning, we did find that participants who gained awareness of the rule showed significantly higher accuracy on nonword trials compared to participants who remained unaware (see Supporting Information). However, even participants classified as "rule-aware" based on their responses on the post-experiment interview did not necessarily consistently use the animacy rule to guide their decisions for nonword categorization. Overall, performance on nonword trials was low overall and variable at the individual level, and thus ultimately could not serve as a reliable measure of an individual's awareness of the rule.

We also found low rates of rule awareness as assessed by verbal reports during the post-experiment interview, with only 38% of children and 23% of adults reporting explicit knowledge of at least one of the four articles. This is fewer than in the previous Batterink et al. (2014) study, in which about half of the adult participants became aware of the hidden rule. Our lower rates of rule awareness may be due to task differences, as participants in our study made concurrent (rather than sequential) animacy and distance judgements, possibly giving them less opportunity to consider animacy as a potentially relevant, isolated factor. An additional possible factor is that the current paradigm included violation trials as well as nonword trials, such that canonical trials represented only ~70% of the total number of trials (rather than ~86% in Batterink et al., 2014). This may have made it more difficult for participants to extract the probabilistic hidden animacy rule, preventing many participants from becoming aware of the rule (and from successfully generalizing the rule) during the training period. More broadly, these findings highlight that artificial language paradigms, including the current task, are generally designed to model one particular aspect of language learning in isolation and over a short period of time. A clear and comprehensive picture of children's language learning ability will emerge by taking into account both laboratory learning approaches as well as ecological studies on age of acquisition effects in second language learners (as reviewed in the Introduction).

A potential limitation of our design is that only children were monitored by an experimenter over Zoom during the experimental task, while adults completed the task without monitoring. We reasoned that some children may have difficulty progressing through the somewhat long and repetitive task and may benefit from experimenter guidance, whereas adults should be easily able to complete the task independently. By monitoring the children via Zoom, we hoped to improve their performance for the assigned task, better equating performance for the explicitly instructed rule between the two groups. While we consider it unlikely that this difference could contribute to our main findings regarding age differences in hidden rule learning, future studies should consider keeping this aspect of the procedure the same between the two groups.

4.2 | No evidence for retention advantages or differences in children

Turning to our retention-related analyses, both children and adults showed a delay-related *decrease* in sensitivity to the hidden animacy rule, as revealed by comparing the BIS violation effect for block 2_A to the other three blocks. Notably, children showed a significant violation effect in the first two blocks, but not in the block immediately following the delay period, such that there is evidence in children (but not in adults) for a true transient destabilization of existing hidden rule knowledge over the 12-h delay. In contrast, for the explicit distance rule, we found that children showed a delay-related *boost* in performance, whereas adults showed relatively stable performance over the delay period, with no significant delay-related change.

The pattern of results shown by children in block 2_A are especially striking (Figure 3), demonstrating a clear boost in explicit distance rule performance alongside a transient decrease in implicit sensitivity to the hidden animacy rule. Though our design does not allow for testing the role of sleep directly, these results align with the idea that children's uniquely rich slow-wave sleep may preferentially consolidate explicit or declarative memories, while preventing consolidation of implicit or procedural memory, as typically seen in adults (Fischer et al., 2007; Giganti et al., 2014; Henderson et al., 2012; Wilhelm et al., 2008; Wilhelm, Prehn-Kristensen, Born, 2012). This memory consolidation trade-off in children has been demonstrated previously using motor-based tasks such as the serial reaction time task, but to our knowledge has not been previously reported within a language learning context. Given that children appear to rapidly discover linguistic patterns during real-time language exposure, without requiring any period of offline consolidation (as evidenced by our rate-based analyses), it is likely that children's overall language development benefits from this trade-off, whereby priority for offline consolidation is given to declarative aspects of language, such as vocabulary.

In adults, we found no significant evidence of an overall gain in sensitivity to the hidden rule across the delay period. This result was somewhat unexpected, as in our previous nap study, we found that adult nappers with longer durations of both slow-wave and rapid eye movement sleep showed an increase in sensitivity to the hidden rule over the nap period (Batterink et al., 2014). However, adults in the current study did not show reliable sensitivity to the animacy rule until the second half of the second session (block $2_{B;}$ Figure 3). Thus, one possibility is that adults' learning of animacy rule prior to the delay period was too weak or unstable to benefit from sleep-related consolidation. This explanation follows from Stickgold's (Stickgold, 2009) theory that the extent of memory consolidation depends on its initial strength, and follows an inverted U-shaped curve, where intermediate levels of performance show the greatest benefit from sleep-dependent consolidation (for additional evidence for this idea, see Cairney et al., 2016; Creery et al., 2015; Wilhelm, Metzkow-Mészàros et al., 2012).

Finally, based on some prior findings (Wilhelm et al., 2013), we had originally hypothesized that children may show a stronger likelihood of becoming explicitly aware of the underlying rule after the delay period, relative to adults. However, we did not find evidence to support this hypothesis; although children did show better generalization performance than adults (Figure 4), this advantage was not significantly different between the first and second session. As described previously, low overall generalization performance precluded us from being able to use performance on nonword trials as a good measure of rule awareness.

4.3 | Conclusion

While children are disadvantaged compared to adults on most high-level cognitive tasks, there are several domains—including ultimate attainment of grammatical rules—where children have been found to outperform adults (Gualtieri & Finn, 2022). Here, we found that children rapidly gained sensitivity to a novel, hidden grammatical rule and were able to generalize this rule to novel words without semantic meaning. In contrast, adults showed reliable sensitivity to the rule only after extended exposure to the artificial article system and a period of consolidation. These findings provide insight into the nature of the critical period for language acquisition, supporting the view that children have an advantage in their ability to implicitly acquire linguistic rules during real-time language exposure. We found no evidence that offline consolidation processes were responsible for this advantage. By demonstrating that children learned more rapidly than adults under equivalent learning conditions, our results indicate that children's long-term superiority for language cannot be driven solely by social and environmental factors, but to differences that are intrinsic to younger learners. Children's intrinsic advantage for language learning is likely to stem both from their (reduced) prior experience as well as maturational factors (Siegelman et al., 2018; Thiessen et al., 2016; Werker & Hensch, 2015).

ACKNOWLEDGMENTS

We are grateful to Jonathan Shammas for his assistance in running participants. This work was funded by Natural Sciences and Engineering Research Council of Canada (RGPIN-2019-05132).

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

DATA AVAILABILITY STATEMENT

The data, analysis code and materials for this study are publicly accessible at https://osf.io/p7h5u/

ORCID

Sarah Berger https://orcid.org/0000-0002-5385-7812
Laura J. Batterink https://orcid.org/0000-0002-6879-4666

ENDNOTE

[1] However, we note that during the course of implicit learning, explicit or consciously-accessible knowledge may develop in parallel. The development of conscious knowledge is optional, may be sporadic, and can often be dissociated experimentally from implicit knowledge (e.g., Batterink et al., 2015; Esser et al., 2022; Higham, 1997; Lieberman et al., 2004; Meulemans & Van der Linden, 1997; Vokey & Brooks, 1992; Willingham & Goedert-Eschmann, 1999; Willingham et al., 2002).

REFERENCES

Adi-Japha, E., Badir, R., Dorfberger, S., & Karni, A. (2014). A matter of time: Rapid motor memory stabilization in childhood. *Developmental Science*, *17*(3), 424–433. https://doi.org/10.1111/desc.12132

Asher, J. J., & Price, B. S. (1967). The learning strategy of the total physical response: Some age differences. *Child Development*, *38*(4), 1219–1227. https://doi.org/10.2307/1127119

Ashtamker, L., & Karni, A. (2013). Motor memory in childhood: Early expression of consolidation phase gains. *Neurobiology of Learning and Memory*, *106*, 26–30. https://doi.org/10.1016/j.nlm.2013.07.003

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Batterink, L. J., Oudiette, D., Reber, P. J., & Paller, K. A. (2014). Sleep facilitates learning a new linguistic rule. *Neuropsychologia*, *65*, 169–179. https://doi.org/10.1016/j.neuropsychologia.2014.10.024

Batterink, L. J., Reber, P. J., Neville, H. J., & Paller, K. A. (2015). Implicit and explicit contributions to statistical learning. *Journal of Memory and Language*, *83*, 62–78. https://doi.org/10.1016/j.jml.2015.04.004

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, *44*(2), 235–249. https://doi.org/10.1006/jmla.2000.2750

Bishop, D. V. M., Barry, J. G., & Hardiman, M. J. (2012). Delayed retention of new word-forms is better in children than adults regardless of language ability: A factorial two-way study. *PLoS ONE*, *7*(5), e37326. https://doi.org/10.1371/journal.pone.0037326

Blanco, N. J., & Sloutsky, V. M. (2019). Adaptive flexibility in category learning? Young children exhibit smaller costs of selective attention than adults. *Developmental Psychology*, *55*(10), 2060–2076. https://doi.org/10.1037/dev0000777

Brehm, L., & Alday, P. M. (2022). Contrast coding choices in a decade of mixed models. *Journal of Memory and Language*, *125*, 104334. https://doi.org/10.1016/j.jml.2022.104334

Cairney, S. A., Lindsay, S., Sobczak, J. M., Paller, K. A., & Gaskell, M. G. (2016). The benefits of targeted memory reactivation for consolidation in sleep are contingent on memory accuracy and direct cue-memory associations. *Sleep*, *39*(5), 1139–1150. https://doi.org/10.5665/sleep.5772

Creery, J. D., Oudiette, D., Antony, J. W., & Paller, K. A. (2015). Targeted memory reactivation during sleep depends on prior learning. *Sleep*, *38*(5), 755–763. https://doi.org/10.5665/sleep.4670

D'Anselmo, A., Prete, G., Zdybek, P., Tommasi, L., & Brancucci, A. (2019). Guessing meaning from word sounds of unfamiliar languages: A cross-cultural sound symbolism study. *Frontiers in Psychology*, *10*, 593. https://doi.org/10.3389/fpsyg.2019.00593

Decker, J. H., Lourenco, F. S., Doll, B. B., & Hartley, C. A. (2015). Experiential reward learning outweighs instruction prior to adulthood. *Cognitive, Affective & Behavioral Neuroscience*, *15*(2), 310–320. https://doi.org/10.3758/s13415-014-0332-5

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, *22*(4), 499–533. https://doi.org/10.1017/S0272263100004022

Dorfberger, S., Adi-Japha, E., & Karni, A. (2007). Reduced susceptibility to interference in the consolidation of motor memory before adolescence. *PLOS ONE*, *2*(2), e240. https://doi.org/10.1371/journal.pone.0000240

Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, *48*(1), 71–99. https://doi.org/10.1016/0010-0277(93)90058-4

Esser, S., Lustig, C., & Haider, H. (2022). What triggers explicit awareness in implicit sequence learning? Implications from theories of consciousness. *Psychological Research*, *86*(5), 1442–1457. https://doi.org/10.1007/s00426-021-01594-3

Ferman, S., & Karni, A. (2010). No childhood advantage in the acquisition of skill in using an artificial language rule. *PLoS ONE*, *5*(10), e13648. https://doi.org/10.1371/journal.pone.0013648

Filoteo, J. V., Lauritzen, S., & Maddox, W. T. (2010). Removing the frontal lobes: The effects of engaging executive functions on perceptual category learning. *Psychological Science*, *21*(3), 415–423. https://doi.org/10.1177/0956797610362646

Fischer, S., Wilhelm, I., & Born, J. (2007). Developmental differences in sleep's role for implicit off-line learning: Comparing children with adults. *Journal of Cognitive Neuroscience*, *19*(2), 214–227. https://doi.org/10.1162/jocn.2007.19.2.214

Flege, J., & Liu, S. (2001). The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, *23*(4), 527–552. https://doi.org/10.1017/S0272263101004041

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language*, *41*(1), 78–104. https://doi.org/10.1006/jmla.1999.2638

Fletcher, P. C., Zafiris, O., Frith, C. D., Honey, R. A. E., Corlett, P. R., Zilles, K., & Fink, G. R. (2005). On the benefits of not trying: Brain activity and connectivity reflecting the interactions of explicit and implicit sequence learning. *Cerebral Cortex*, *15*(7), 1002–1015. https://doi.org/10.1093/cercor/bhh201

Foerde, K., Knowlton, B. J., & Poldrack, R. A. (2006). Modulation of competing memory systems by distraction. *Proceedings of the National Academy of Sciences*, *103*(31), 11778–11783. https://doi.org/10.1073/pnas.0602659103

Frank, S. M., Becker, M., Qi, A., Geiger, P., Frank, U. I., Rosedahl, L. A., Malloni, W. M., Sasaki, Y., Greenlee, M. W., & Watanabe, T. (2022). Efficient learning in children with rapid GABA boosting during and after training. *Current Biology*, *32*(23), 5022–5030.e7. https://doi.org/10.1016/j.cub.2022.10.021

Frensch, P. A., & Rünger, D. (2003). Implicit learning. *Current Directions in Psychological Science*, *12*(1), 13–18. https://doi.org/10.1111/1467-8721.01213

Gaskell, M. G., Warker, J., Lindsay, S., Frost, R., Guest, J., Snowdon, R., & Stackhouse, A. (2014). Sleep Underpins the Plasticity of Language Production. *Psychological Science*, *25*(7), 1457–1465.

Giganti, F., Arzilli, C., Conte, F., Toselli, M., Viggiano, M. P., & Ficca, G. (2014). The effect of a daytime nap on priming and recognition tasks in preschool children. *Sleep*, *37*(6), 1087–1093. https://doi.org/10.5665/sleep.3766

Gualtieri, S., & Finn, A. S. (2022). The sweet spot: When children's developing abilities, brains, and knowledge make them better learners than adults. *Perspectives on Psychological Science*, *17*(5), 1322–1338. https://doi.org/10.1177/17456916211045971

Hakuta, K., Bialystok, E., & Wiley, E. (2003). Critical evidence: A test of the critical-period hypothesis for second-language acquisition. *Psychological Science*, *14*(1), 31–38. https://doi.org/10.1111/1467-9280.01415

Hartshorne, J. K., Tenenbaum, J. B., & Pinker, S. (2018). A critical period for second language acquisition: Evidence from 2/3 million English speakers. *Cognition*, *177*, 263–277. https://doi.org/10.1016/j.cognition.2018.04.007

Henderson, L., Weighall, A., Brown, H., & Gaskell, G. (2012). Consolidation of vocabulary is associated with sleep in children. *Developmental Science*, *15*(5), 674–687. https://doi.org/10.1111/j.1467-7687.2012.01172.x

Higham, P. A. (1997). Dissociations of grammaticality and specific similarity effects in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1029–1045. https://doi.org/10.1037/0278-7393.23.4.1029

Hudson Kam, C. L., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, *59*(1), 30–66. https://doi.org/10.1016/j.cogpsych.2009.01.001

Johnson, J., & Newport, E. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, *21*(1), 60–99. https://doi.org/10.1016/0010-0285(89)90003-0

Jung, Y., Forest, T. A., Walther, D. B., & Finn, A. S. (2023). Neither enhanced nor lost: The unique role of attention in children's neural representations. *Journal of Neuroscience*, *43*(21), 3849–3859. https://doi.org/10.1523/JNEUROSCI.0159-23.2023

Kareev, Y. (1995). Through a narrow window: Working memory capacity and the detection of covariation. *Cognition*, *56*(3), 263–269. https://doi.org/10.1016/0010-0277(95)92814-G

Korman, M., Doyon, J., Doljansky, J., Carrier, J., Dagan, Y., & Karni, A. (2007). Daytime sleep condenses the time course of motor memory consolidation. *Nature Neuroscience*, *10*(9), 1206–1213. https://doi.org/10.1038/nn1959

Krashen, S. D., Long, M. A., & Scarcella, R. C. (1979). Age, rate and eventual attainment in second language acquisition. *TESOL Quarterly*, *13*(4), 573. https://doi.org/10.2307/3586451

Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990. https://doi.org/10.3758/s13428-012-0210-4

Lenneberg, E. H., Chomsky, N., & Marx, O. (1967). *Biological foundations of language* (99th ed.). John Wiley & Sons Inc.

Lenth, R., Beurkner, P., Herve, M., Love, J., Miguez, F., Riebl, H., & Singmann, H. (2022). Emmeans: Estimated marginal means, aka least-squares Means (1.7.3). https://cran.r-project.org/web/packages/emmeans/index.html

Lichtman, K. (2016). Age and learning environment: Are children implicit second language learners?*. *Journal of Child Language*, *43*(3), 707–730. https://doi.org/10.1017/S0305000915000598

Lieberman, M. D., Chang, G. Y., Chiao, J., Bookheimer, S. Y., & Knowlton, B. J. (2004). An Event-related fMRI Study of Artificial Grammar Learning in a Balanced Chunk Strength Design. *Journal of Cognitive Neuroscience*, *16*(3), 427–438. https://doi.org/10.1162/089892904322926764

Liesefeld, H. R., Fu, X., & Zimmer, H. D. (2015). Fast and careless or careful and slow? Apparent holistic processing in mental rotation is explained by speed-accuracy trade-offs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(4), 1140–1151. https://doi.org/10.1037/xlm0000081

Liesefeld, H. R., & Janczyk, M. (2019). Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behavior Research Methods*, *51*(1), 40–60. https://doi.org/10.3758/s13428-018-1076-x

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, *12*(3), 251–285. https://doi.org/10.1017/S0272263100009165

Marinova-Todd, S. H., Marshall, D. B., & Snow, C. E. (2000). Three misconceptions about age and L2 learning. *TESOL Quarterly*, *34*(1), 9–34. https://doi.org/10.2307/3588095

Mayberry, R. I. (1993). First-language acquisition after childhood differs from second-language acquisition. *Journal of Speech, Language, and Hearing Research*, *36*(6), 1258–1270. https://doi.org/10.1044/jshr.3606.1258

Meulemans, T., & Van der Linden, M. (1997). Associative chunk strength in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 1007–1028. https://doi.org/10.1037/0278-7393.23.4.1007

Nemeth, D., Janacsek, K., Polner, B., & Kovacs, Z. A. (2013). Boosting human learning by hypnosis. *Cerebral Cortex*, *23*(4), 801–805. https://doi.org/10.1093/cercor/bhs068

Newport, E. L. (1988). Constraints on learning and their role in language acquisition: Studies of the acquisition of American sign language. *Language Sciences*, *10*(1), 147–172. https://doi.org/10.1016/0388-0001(88)90010-1

Newport, E. L. (1990). Maturational constraints on language learning. *Cognitive Science*, *14*(1), 11–28. https://doi.org/10.1207/s15516709cog1401_2

Newport, E. L., Bavelier, D., & Neville, H. J. (2001). Critical thinking about critical periods: Perspectives on a critical period for language acquisition. In E. Dupoux (Ed.), *Language, brain, and cognitive development: Essays in honor* of Jacques Mehler (pp. 481–502). The MIT Press.

Nissen, M., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures*1. *Cognitive Psychology*, *19*(1), 1–32. https://doi.org/10.1016/0010-0285(87)90002-8

Paradis, M. (2009). *Declarative and procedural determinants of second languages*. John Benjamins Publishing Company. https://benjamins.com/catalog/sibil.40

Peiffer, A., Brichet, M., De Tiège, X., Peigneux, P., & Urbain, C. (2020). The power of children's sleep—Improved declarative memory consolidation in children compared with adults. *Scientific Reports*, *10*(1), 9979. https://doi.org/10.1038/s41598-020-66880-3

Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior research methods*, *51*(1), 195–203. https://doi.org/10.3758/s13428-018-01193-y

Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences*, *10*(5), 233–238. https://doi.org/10.1016/j.tics.2006.03.006

Plebanek, D. J., & Sloutsky, V. M. (2017). Costs of selective attention: When children notice what adults miss. *Psychological Science*, *28*(6), 723–732. https://doi.org/10.1177/0956797617693005

Poldrack, R. A., Clark, J., Paré-Blagoev, E. J., Shohamy, D., Creso Moyano, J., Myers, C., & Gluck, M. A. (2001). Interactive memory systems in the human brain. *Nature*, *414*(6863), 546–550. Article 6863. https://doi.org/10.1038/35107080

Prehn-Kristensen, A., Göder, R., Chirobeja, S., Bressmann, I., Ferstl, R., & Baving, L. (2009). Sleep in children enhances preferentially emotional declarative but not procedural memories. *Journal of Experimental Child Psychology*, *104*(1), 132–139. https://doi.org/10.1016/j.jecp.2009.01.005

Ramscar, M., & Gitcho, N. (2007). Developmental change and the nature of learning in childhood. *Trends in Cognitive Sciences*, *11*(7), 274–279. https://doi.org/10.1016/j.tics.2007.05.007

Rastle, K., Harrington, J., & Coltheart, M. (2002). 358,534 nonwords: The ARC Nonword Database. *The Quarterly Journal of Experimental Psychology Section A*, *55*(4), 1339–1362. https://doi.org/10.1080/02724980244000099

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, *118*, 219–235. https://doi.org/10.1037/0096-3445.118.3.219

Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition*, *177*, 198–213. https://doi.org/10.1016/j.cognition.2018.04.011

Smalle, E. H. M., Daikoku, T., Szmalec, A., Duyck, W., & Möttönen, R. (2022). Unlocking adults' implicit statistical learning by cognitive depletion. *Proceedings of the National Academy of Sciences*, *119*(2), e2026011119. https://doi.org/10.1073/pnas.2026011119

Smalle, E. H. M., Muylle, M., Szmalec, A., & Duyck, W. (2017). The different time course of phonotactic constraint learning in children and adults: Evidence from speech errors. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *43*(11), 1821–1827. https://doi.org/10.1037/xlm0000405

Smalle, E. H. M., Page, M. P. A., Duyck, W., Edwards, M., & Szmalec, A. (2018). Children retain implicitly learned phonological sequences better than adults: A longitudinal study. *Developmental Science*, *21*(5), e12634. https://doi.org/10.1111/desc.12634

Snow, C. E., & Hoefnagel-Höhle, M. (1978). The critical period for language acquisition: Evidence from second language learning. *Child Development*, *49*(4), 1114–1128. https://doi.org/10.2307/1128751

Stickgold, R. (2009). How Do I Remember? Let Me Count the Ways. *Sleep Medicine Reviews*, *13*(5), 305–308. https://doi.org/10.1016/j.smrv.2009.05.004

Thiessen, E. D., Girard, S., & Erickson, L. C. (2016). Statistical learning and the critical period: How a continuous learning mechanism can give rise to discontinuous learning: Statistical learning and the critical period. *Wiley Interdisciplinary Reviews: Cognitive Science*, *7*(4), 276–288. https://doi.org/10.1002/wcs.1394

Thompson-Schill, S. L., Ramscar, M., & Chrysikou, E. G. (2009). Cognition without control: When a little frontal lobe goes a long way. *Current Directions in Psychological Science*, *18*(5), 259–263. https://doi.org/10.1111/j.1467-8721.2009.01648.x

Ullman, M. T. (2001). A neurocognitive perspective on language: The declarative/procedural model. *Nature Reviews Neuroscience*, *2*(10), 717–726. https://doi.org/10.1038/35094573

Ullman, M. T. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition*, *92*(1–2), 231–270. https://doi.org/10.1016/j.cognition.2003.10.008

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice*, (pp. 141–178). Georgetown University Press.

Vokey, J. R., & Brooks, L. R. (1992). Salience of item knowledge in learning artificial grammars. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 328–344. https://doi.org/10.1037/0278-7393.18.2.328

Wang, J.-Y., Weber, F. D., Zinke, K., Inostroza, M., & Born, J. (2018). More effective consolidation of episodic long-term memory in children than adults-unrelated to sleep. *Child Development*, *89*(5), 1720–1734. https://doi.org/10.1111/cdev.12839

Warker, J. A. (2013). Investigating the Retention and Time-Course of Phonotactic Constraint Learning From Production Experience. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *39*(1), https://doi.org/10.1037/a0028648

Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, *8*(3), 231–256. https://doi.org/10.1162/jocn.1996.8.3.231

Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: New directions. *Annual Review of Psychology*, *66*(1), 173–196. https://doi.org/10.1146/annurev-psych-010814-015104

Wilhelm, I., Diekelmann, S., & Born, J. (2008). Sleep in children improves memory performance on declarative but not procedural tasks. *Learning & Memory*, *15*(5), 373–377. https://doi.org/10.1101/lm.803708

Wilhelm, I., Metzkow-Mészàros, M., Knapp, S., & Born, J. (2012). Sleep-dependent consolidation of procedural motor memories in children and adults: The pre-sleep level of performance matters. *Developmental Science*, *15*(4), 506–515. https://doi.org/10.1111/j.1467-7687.2012.01146.x

Wilhelm, I., Prehn-Kristensen, A., & Born, J. (2012). Sleep-dependent memory consolidation—What can be learnt from children? *Neuroscience & Biobehavioral Reviews*, *36*(7), 1718–1728. https://doi.org/10.1016/j.neubiorev.2012.03.002

Wilhelm, I., Rose, M., Imhof, K. I., Rasch, B., Büchel, C., & Born, J. (2013). The sleeping child outplays the adult's capacity to convert implicit into explicit knowledge. *Nature Neuroscience*, *16*(4), 391–393. https://doi.org/10.1038/nn.3343

Williams, J. N. (2005). Learning without awareness. Studies in second language acquisition (Vol. *27*(2), pp. 269–304). Cambridge University Press. https://doi.org/10.1017/S0272263105050138

Willingham, D. B., & Goedert-Eschmann, K. (1999). The Relation Between Implicit and Explicit Learning: Evidence for Parallel Development. *Psychological Science*, *10*(6), 531–534. https://doi.org/10.1111/1467-9280.00201

Willingham, D. B., Salidis, J., & Gabrieli, J. D. E. (2002). Direct Comparison of Neural Systems Mediating Conscious and Unconscious Skill Learning. *Journal of Neurophysiology*, *88*(3), 1451–1460. https://doi.org/10.1152/jn.2002.88.3.1451

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Berger, S., & Batterink, L. J. (2024). Children extract a new linguistic rule more quickly than adults. *Developmental Science*, e13498. https://doi.org/10.1111/desc.13498