



Positional encoding in cotton-top tamarins (*Saguinus oedipus*)

Elisabetta Versace^{1,2} · Jessica R. Rogge² · Natalie Shelton-May² · Andrea Ravignani^{3,4}

Received: 3 March 2019 / Revised: 29 May 2019 / Accepted: 5 June 2019 / Published online: 1 July 2019
© The Author(s) 2019

Abstract

Strategies used in artificial grammar learning can shed light into the abilities of different species to extract regularities from the environment. In the $A(X)^nB$ rule, A and B items are linked, but assigned to different positional categories and separated by distractor items. Open questions are how widespread is the ability to extract positional regularities from $A(X)^nB$ patterns, which strategies are used to encode positional regularities and whether individuals exhibit preferences for absolute or relative position encoding. We used visual arrays to investigate whether cotton-top tamarins (*Saguinus oedipus*) can learn this rule and which strategies they use. After training on a subset of exemplars, two of the tested monkeys successfully generalized to novel combinations. These tamarins discriminated between categories of tokens with different properties (A , B , X) and detected a positional relationship between non-adjacent items even in the presence of novel distractors. The pattern of errors revealed that successful subjects used visual similarity with training stimuli to solve the task and that successful tamarins extracted the relative position of A s and B s rather than their absolute position, similarly to what has been observed in other species. Relative position encoding appears to be favoured in different tasks and taxa. Generalization, though, was incomplete, since we observed a failure with items that during training had always been presented in reinforced arrays, showing the limitations in grasping the underlying positional rule. These results suggest the use of local strategies in the extraction of positional rules in cotton-top tamarins.

Keywords Rule learning · Cotton-top tamarins · Positional rule · Relative position · Absolute position · Artificial grammar learning · Non-adjacent dependency

Introduction

Extracting regularities is necessary to make sense of the numerous stimuli available in the environment. The relative location of different items in time and space is important in domains as different as causal reasoning (A hit B vs B hit

A), spatial navigation (A to the left of B), language (“A hit B” vs “hit B A” vs “B hit A”), and action planning (“grasp A, then pierce B”). Comparing strategies and constraints in learning positional regularities across species is a way to understand cognitive species specificities and shared abilities to process environmental regularities (Chen et al. 2014; Fitch 2017; Fitch and Friederici 2012; Ghirlanda et al. 2017; Stobbe et al. 2012).

Artificial grammar paradigms have been used to investigate abilities relevant for language processing. In this context, the investigation of positional categories (e.g., the A category of items located in first position vs the B category of items located in second position or A items located to the left of B items) has a long history. Smith (1966) exposed human subjects to four sets of letter pairs built from four classes of letters—M, N, P, and Q—that formed MN and PQ sequences. When tested in free recall, subjects produced more intrusions (wrong recalled items) of the form MQ and PN and less MP and QN intrusions than expected by chance. Along this line, Reber and Lewis (1977) showed that letters

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10071-019-01277-y>) contains supplementary material, which is available to authorized users.

✉ Elisabetta Versace
e.versace@qmul.ac.uk

- ¹ Department of Biological and Experimental Psychology, Queen Mary University of London, London E1 4NS, UK
- ² Department of Psychology, Harvard University, Cambridge, USA
- ³ Artificial Intelligence Lab, Vrije Universiteit Brussel, 1050 Brussels, Belgium
- ⁴ Language and Cognition Department, Max Planck Institute for Psycholinguistics, 6525 XD Nijmegen, The Netherlands

presented in the initial and terminal position of a string are particularly salient. These results are in line with a positional encoding of the items. Some questions remain open: (i) to which extent the capacity to encode positional regularities between categories of items is widespread among different species and (ii) whether the relative or absolute position of the items is preferentially encoded. Relative position encoding concerns the position of an item with respect to others, irrespective of the position within the pattern, while absolute position encoding takes into account the ordinal position from at least one of the edges or the absolute distance from the edges.

Some studies have investigated the capabilities of positional rule learning in non-human animals. Rats can use the serial/temporal order of sequentially presented elements to discriminate among them. In particular, rats can encode the sequential structure of two elements, as in $A \rightarrow B$ (Murphy et al. (2004), or three elements, as in XYX vs XXY and YXX (Murphy et al. (2008). Using long-distance dependencies (relations between non-adjacent items), Endress et al. (2010) found that violations at the edges are more salient for chimpanzees than violations within the sequence of items. These results indicate a possible specialization in processing items located at the edges. The aforementioned studies presented stimuli sequences showing tokens sequentially, thus introducing memory and attention requirements that could interfere with the performance and comparison of computational capacities across species (Fitch 2014; Fitch et al. 2012; Frank and Gibson 2011). This issue can be overcome using visual arrays, with patterns defined by spatial relationships and all the relevant components presented at the same time. Fiser and Aslin (2001, 2002, 2005) showed that after a familiarization phase, human subjects can use the frequency of occurrence of single shapes, the shape position in an array, and the arrangement of shape pairs to discriminate between familiar and unfamiliar configurations. These results suggest that computations operating on serial stimuli can be available also for visual configurations. Along this line, Rosa-Salva et al. (2018) have tested the spontaneous abilities of domestic chicks in visual imprinting, showing a preferential encoding for the position of items located at the edges. The similarities and differences in processing visual and acoustic regularities have just started to be explored (review in Milne et al. 2018).

In the visual modality, the capacity to process rules using visual arrays has been investigated in different species (Grainger et al. 2012; Murphy et al. 2008; Ravignani and Sonnweber 2017; Rey et al. 2012; Scarf et al. 2016; Sonnweber et al. 2015; Stobbe et al. 2012; Versace et al. 2006, 2017), but the processing of positional regularities has just started to be clarified. Here, we use visual arrays to test the ability of a small non-human primate, the cotton-top tamarin (*Saguinus oedipus*), to encode the positional regularity

$A(X)^nB$ (Chomsky 1956). According to this grammar, A and B items are linked but assigned to different positional categories, with A s located to the left of B s. The A and B categories exhibit the positional regularities of the first and last tokens of the MN–PQ grammar introduced by Smith (1966). We used three category A tokens— A_1 , A_2 , and A_3 , four category B tokens— B_1 , B_2 , B_3 , and B_4 , and several X distractor tokens (Fig. 1). The full mastery of this positional grammar requires subjects to treat familiar and unfamiliar $A(X)^nB$ combinations as grammatical and familiar or unfamiliar $B(X)^nA$, $A(X)^nA$, and $B(X)^nB$ tokens as ungrammatical. If subjects master the grammar, they should treat configurations of $A(X)^nB$ that they have not previously experienced as grammatical. Failure with specific configurations would instead support an encoding based on local features, instead of the overall grammar. We first trained tamarins on a subset of configurations, reinforcing choices of stimuli consistent with the $A(X)^nB$ pattern, and giving no reinforcement for

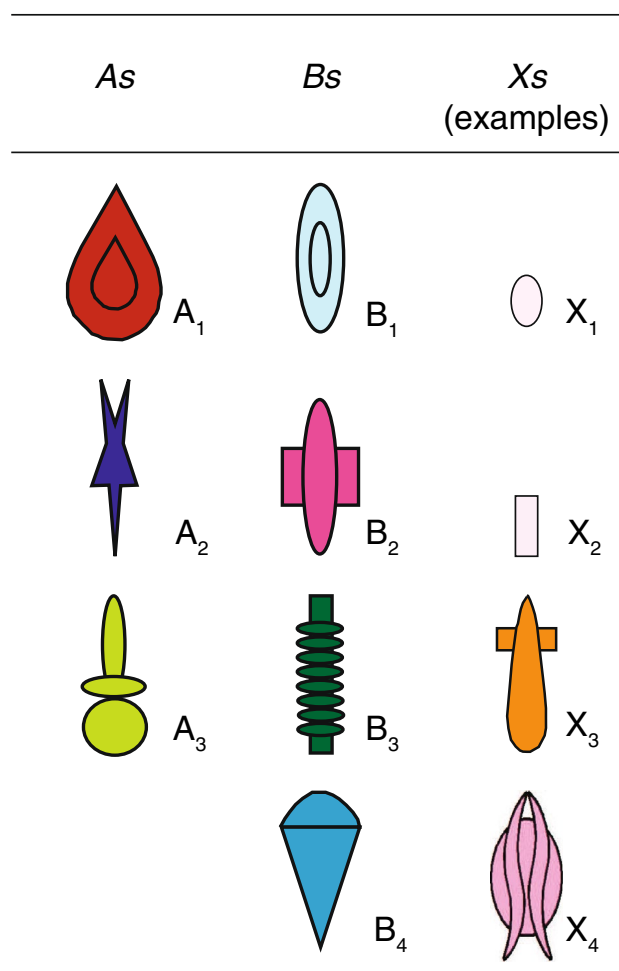


Fig. 1 All A and B tokens used in the experiment and four examples of X tokens. Arrays consistent with the rule have A tokens located to the left of B tokens. Distractor X tokens have no relevance in determining the consistency of an array with the grammar

choices of inconsistent stimuli. Then, we tested tamarins presenting novel items consistent and inconsistent with the rule.

$A(X)^nB$ configurations are compatible with both absolute position rules (“A located on the left edge”) and relative position rules (“A to the left of B”). We used test items with distractors located at the edges (e.g., $XAXXBX$ vs $XBXXAX$) to clarify whether an individual trained on recognizing $A(X)^nB$ configurations had extracted the absolute or relative position of As and Bs with respect to the edges. Moreover, by varying the number of central distractors, we probed the relative/absolute encoding with respect to the center of the array.

General methods

The experimental schedule alternated training and test sessions. We first trained tamarins on a subset of stimuli consistent or inconsistent with the $A(X)^nB$ rule. This rule specifies the relative position of A and B tokens. We used three category A tokens ($A_1, A_2,$ and A_3), four category B tokens ($B_1, B_2, B_3,$ and B_4), and a total of 17 category X tokens (Fig. 1). A and B tokens were assigned to these categories in such a way that no obvious feature could be used to distinguish the two categories (e.g., both categories contained items with round and sharp edges, different hues, and luminance).

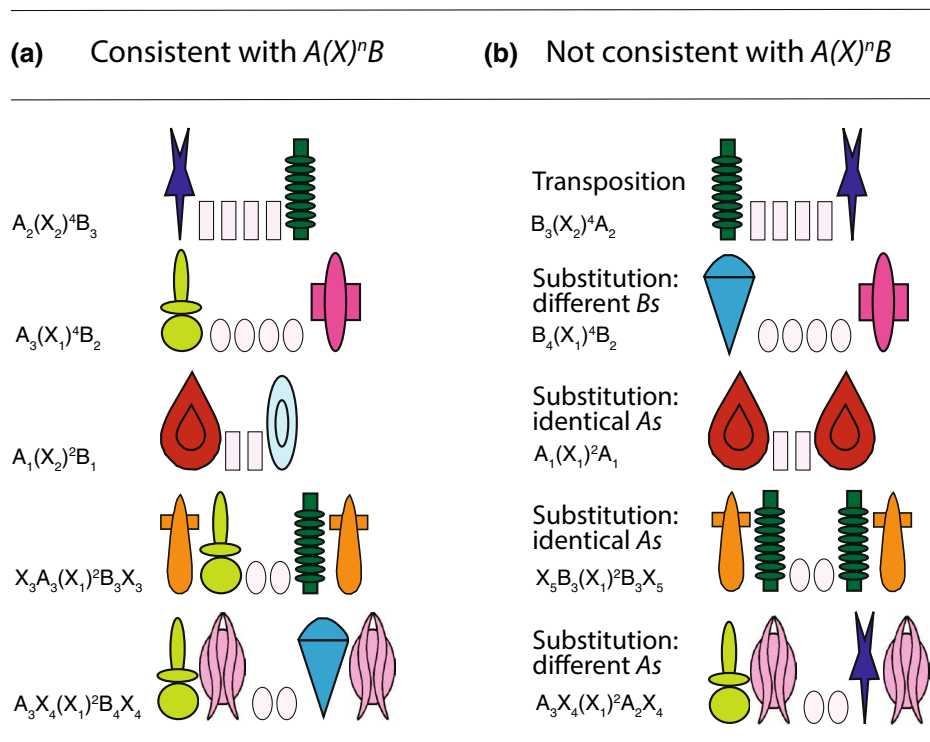
According to the $A(X)^nB$ rule: (i) A tokens must be presented in left position with respect to B tokens; (ii) B tokens

must be presented to the right of A tokens; and (iii) X tokens can vary in number and their position is irrelevant to define the grammaticality of the configuration. When located between As and Bs, X tokens allow us to investigate non-adjacent relationships between As and Bs. Figure 2 shows some examples of the arrays used during the experiments, as $A_2(X_2)^4B_3$, $A_3(X_1)^4B_2$, and $A_1(X_1)^2B_1$ for the grammatical arrays (Fig. 2a). In ungrammatical arrays (Fig. 2b), the relative position of A and B tokens within the visual array was swapped, as in $B_3(X_2)^4A_2$ (transposition), or one token was misplaced as in $B_4(X_1)^4B_2$ (substitution, with two identical tokens from the same B category) or $A_1(X_1)^2A_1$ (substitution with two different tokens from the same A category).

During the initial training (training A), we presented subjects with a subset of the possible $A(X)^nB$ combinations and violations of the rule, rewarding only the choices of grammatical stimuli. In the subsequent phases, subjects were tested and trained with new combinations of the tokens used in training A and with new X tokens. In particular, in test 1, we introduced novel arrangements of A and B tokens and in test 2, we presented arrays with two or four new Xs in the middle. Finally, to assess whether each subject used a relative or absolute position strategy to solve the task, in tests 3 and 4, we placed novel Xs in different locations both on the edges and within the array.

Test 1 was designed to test tamarins’ capability to extract and generalize the regularity presented during the training. Yet, succeeding in test 1 does not clarify whether tamarins encoded the relative position of A vs B tokens (“A to the left

Fig. 2 **a** Examples of arrangements consistent with the $A(X)^nB$ rule. **b** Examples of arrangements not consistent with the $A(X)^nB$ rule. In transposition violations, the position of As and Bs is swapped, in substitution violations two tokens of the same category are presented in the same array, so that only one is correctly located with respect to the other. Identity violations show two identical A or B tokens. Distractor tokens could be located on the edges or within the array, thus modifying the absolute position of As and Bs



of B ”) or the absolute position of A and B with respect to the edges (“ A in the extreme left position, B in the extreme right position”). In fact, both the relative and absolute position encoding were consistent with the stimuli presented during the initial training. The strategy used to solve the task, more than the mere success, can inform us about representations, computational processes, and biases of the subjects. Tests 2 through 4 were designed to investigate whether and how tamarins extracted a relative position or an absolute position regularity from the stimuli presented during the training.

Subjects

We tested four adult cotton-top tamarins, one male (RK) and three females (RB, SH, and EM), housed at Harvard University. All subjects were born in captivity and socially housed, with separate home cages for each breeding pair and their offspring. Subjects were maintained on a diet of monkey chow, fruit, seeds, and mealworms, together with free access to water. Subjects voluntarily left their home cages lured out by a piece of raisin.

Apparatus

During the experimental sessions, subjects were individually housed in a wire mesh box ($31 \times 31 \times 25$ cm) that contained two pulling tools located under laminated plastic cards. The cards presented the target (rewarded and unrewarded) stimuli. Monkeys could access the two pulling tools presented on an acrylic apparatus ($40 \times 50 \times 6$ cm) through two small holes. Each tool consisted of a pulling stem and a card covering a tray at the end of the stem. When subjects pulled one of the stems the tool advanced, the card flipped back, presenting either the food reward (a small piece of a Froot

Loop© cereal) or nothing at all. Stimuli were presented on a plastic laminated sheet (11.5×7.5 cm) with different, linearly arranged shapes corresponding to the consistent or inconsistent arrays.

Stimuli

Tokens used to compose visual arrays belonged to three categories (A , B , and X). Each category contained distinctive tokens, as shown in Fig. 1. Tokens were arranged in visual arrays, printed on cards, and located on the apparatus. The tokens used and their position within the array determined whether the stimulus printed on each card was consistent or inconsistent with the target rule.

Consistent arrays followed the $A(X)^nB$ rule, for example, $A_1X_1X_1B_2$. In arrays not consistent with the target rule, the position of A and B tokens was swapped, as in $B_2X_1X_1A_1$, or either A or B tokens were not used. X tokens, irrelevant in determining the consistency of the stimuli with the target rule, could vary in size and number, extending or reducing the distance of the dependency between other tokens.

Possible arrangements of the A and B tokens are shown in Table 1: 12 patterns are consistent with the target rule $A(X)^nB$ and 37 are not consistent with it. Arrangements not consistent with the rule include transpositions, in which B s are located to the left of A s, and substitutions, in which two tokens of the same A (or B) category are present. During the initial training, we used only 6 of the grammatical arrangements and 7 of the ungrammatical arrangements, 2 or 4 X_1 and X_2 tokens, saving the other configurations of stimuli for the tests. The 13 patterns employed in the training are shaded grey in Table 1, and shaded in green and red in the electronic version.

Table 1 Possible arrangements of A s and B s

Consistent with $A(X)^nB$	Transposition	Substitution: different A s	Substitution: identical A s	Substitution: different B s	Substitution: identical B s
$A_1(X)^nB_1$	$B_1(X)^nA_1$	$A_1(X)^nA_2$	$A_1(X)^nA_1$	$B_1(X)^nB_2$	$B_1(X)^nB_1$
$A_1(X)^nB_2$	$B_1(X)^nA_2$	$A_1(X)^nA_3$	$A_2(X)^nA_2$	$B_1(X)^nB_3$	$B_2(X)^nB_2$
$A_1(X)^nB_3$	$B_1(X)^nA_3$	$A_2(X)^nA_1$	$A_3(X)^nA_3$	$B_1(X)^nB_4$	$B_3(X)^nB_3$
$A_1(X)^nB_4$	$B_2(X)^nA_1$	$A_2(X)^nA_3$		$B_2(X)^nB_1$	$B_4(X)^nB_4$
$A_2(X)^nB_1$	$B_2(X)^nA_2$	$A_3(X)^nA_1$		$B_2(X)^nB_3$	
$A_2(X)^nB_2$	$B_2(X)^nA_3$	$A_3(X)^nA_2$		$B_2(X)^nB_4$	
$A_2(X)^nB_3$	$B_3(X)^nA_1$			$B_3(X)^nB_1$	
$A_2(X)^nB_4$	$B_3(X)^nA_2$			$B_3(X)^nB_2$	
$A_3(X)^nB_1$	$B_3(X)^nA_3$			$B_3(X)^nB_4$	
$A_3(X)^nB_2$	$B_4(X)^nA_1$				
$A_3(X)^nB_3$	$B_4(X)^nA_2$				
$A_3(X)^nB_4$	$B_4(X)^nA_3$				

Only bold patterns were used during the training. Consistent patterns follow the $A(X)^nB$ rule, inconsistent patterns swap the position of A s and B s (transpositions) or show two different or identical tokens of the same A or B category (substitutions)

Test 1 featured 36 test stimuli consisting of novel arrangements of the tokens previously used during the training (unshaded stimuli in Table 1). In test 2, we introduced four novel *X* tokens in the central position [we used, for example, stimuli such as $A_1(X_3)^nB_2$ and $A_2(X_4)^nA_2$], using 26 new stimuli. In test 3 and test 4, we introduced new *X* tokens in novel positions either at the edges of the arrays [as $X_6A_3(X_1)^2B_3X_7$ and $X_6B_3(X_1)^2A_1X_7$; in test 3, we used 42 new stimuli] or internally [as $A_3X_6(X_1)^2B_4X_7$ and $A_3X_6(X_1)^2X_7A_2$; in test 4, we used 35 new stimuli, see Fig. 2]. *X*s were irrelevant in determining the rule consistency of the stimuli and were used to evaluate the information encoded by tamarins during the training.

Procedure

Before starting the training on the $A(X)^nB$ rule, each subject was familiarized with the apparatus and the experimental procedure. Both training and test sessions involved the following procedure. The target subject was lured out of its home cage with a piece of food into a transport box and then moved individually to the experimental room for an experimental session. Prior to a trial, and out of view from the subject, the experimenter prepared the appropriate stimuli and reward. The overall sequence of different stimulus pairings, along with the right or left position of each card, was randomized and counterbalanced across trials and within sessions. For each session, consistent cards were equally distributed between the right and left sides of the apparatus and no more than two consistent cards were presented on the same side consecutively.

The apparatus was presented for 3 s in a position out of reach from the subject at a distance of approximately 40 cm, and then subsequently pushed towards the subject. In cases where the subject did not look at the setup within 4 s, the experimenter drew the subject’s attention to the tray by pointing to the midpoint between the cards, and then moved the tray forward. The subject was only allowed to pull one of the two tools; following its first selection, the alternative tool was retracted, out of reach. When the subject pulled the tool with the consistent card, a food reward was immediately available. If the subject pulled the tool with the inconsistent card, the experimenter revealed the hidden food under the consistent card. In both training and testing sessions, correct choices were rewarded in the same way. Soon after the subject completed its choice, the apparatus was removed and the experimenter prepared a new trial out of view. If the subject did not make any choice within 8 s, the apparatus was removed and the trial aborted.

Each training session consisted of 2–6 warm-up trials followed by 12 training trials. Warm-up trials continued until two consecutive correct choices were made, at which point the experimenter proceeded the session. If more than

six warm-up trials were necessary, the session was aborted. Warm-up trials consisted of consistent stimuli that the subject had successfully discriminated in the previous conditions (or during the training for the first test), and were designed to make sure that on each session the subject was attentive and motivated; as such, if a subject were attending to the material and motivated to pull the tool, it should succeed on the warm-up trials.

Each test session consisted of 2–6 warm-up trials followed by 12 test trials interspersed with 4 trials with stimuli already presented during the previous training. When responsive, each subject ran two experimental sessions per day. To guarantee an appropriate level of motivation, the inter-session interval within a day was at least 3 h. The difference between training and test sessions was the novel material presented during test sessions.

Monkeys’ responses were coded in terms of which array (consistent or inconsistent) was selected on each trial. To move from the initial training stage (training 1) to the tests, we required subjects to reach a criterion of 40/48 correct trials or better. This corresponds to the cutoff value for a binomial test with $\alpha=0.05$, consisting of 10 out of 12 correct trials over 4 consecutive sessions or better.

To determine whether subjects could discriminate between novel consistent and inconsistent stimuli, showing generalization, we analyzed (a) the scores of the first 48 trials (four sessions) and (b) the scores of the first 96 test trials (eight sessions). We ran the analysis on eight sessions to increase the number of trials and investigate the responses to specific violations.

Experimental schedule

The experimental schedule went through the following stages, as summarized in Table 2: training A, test 1, training b, test 2, test 3, test 4.

Training A lasted until the subject reached the criterion of at least 40/48 correct responses in four consecutive sessions. During training A, we introduced the experimental

Table 2 Experimental schedule, composed by trainings and test stages

Stage	Stimuli
Training A	A subset of 13 out of 49 possible arrangements
Test 1	Novel arrangements
Training B	Stimuli presented in the previous two stages
Test 2	Novel <i>X</i>s in the centre
Test 3	Novel <i>X</i>s on the edges
Training C	Same stimuli presented in Training B
Test 4	Novel <i>X</i>s between <i>As</i>, <i>Bs</i>, and <i>Xs</i>

Bold rows indicate the tests

stimuli. In this stage, subjects were presented with all the *A* and *B* tokens, but only a subset of all the possible combinations of them (see Table 1).

Test 1 explored the extent to which tamarins generalized the distinction between consistent and inconsistent stimuli to new spatial arrangements of the tokens experienced during the training. We hypothesized that if tamarins had encoded the positional regularity of *A* and *B* tokens (“*A* to the left of *B*”) they should, in the absence of further training, choose consistent combinations [i.e., $A(X)^nB$] more often than inconsistent combinations.

Subjects responding above chance to test 1 moved on to training B, in which we presented a total of 36 sessions to each subject, using the same stimuli presented in test 1. Subjects reached the criterion of at least 40/48 correct responses in four consecutive sessions before moving to the next test 2.

Test 2 was identical to test 1, with the exception that experimental trials were composed with four novel *X*s. Test 2 explored whether tamarins could generalize to novel *X*s located in the center of the arrays.

Test 3 was identical to test 1, except that new *X*s were located at the edges of the arrays, so that stimuli followed patterns similar to $X A(X)^nB X$ or $XB(X)^nAX$. Up to this stage, *A* and *B* tokens always occupied the edges of the sequence. Two alternative hypotheses could account for tamarins’ success until test 2. Tamarins could have learned, instead of the relative position of *A* and *B* tokens, their absolute position with respect to the edges. In that case, they would not have been able to generalize to arrays not containing *A* or *B* tokens at the edges, as presented in test 3. As an alternative hypothesis, if tamarins had learned that *A* must be on the left with respect to *B*, their performance should not have been affected by the insertion of novel tokens at the edges. Subjects responding above chance in test 3 proceeded to training C (with the same stimuli used in training B) and then to test 4.

Test 4 was identical to test 1, except that at least one *X* token was located between *A*s or *B*s and the central *X*s, so that stimuli followed patterns of the form $XA(X)^nXB$ or $BX(X)^nAX$. If tamarins learned that the position of *A* and *B* with respect to *X* tokens were irrelevant, or that the symmetry of *A* and *B* with respect to the center of the array were irrelevant, then their performance should not be affected by the insertion of novel *X*s adjacent to the central tokens.

Each test stage was presented 24 times, although only the first 8 sessions were taken into account for statistical analysis, to exclude (or reduce) learning effects (see “Learning during the test and differences between test stages” for an analysis of learning during the tests).

Results and discussion

Training A

Training A lasted until the subject reached the criterion of at least 40/48 correct responses in four consecutive sessions. As expected, subjects varied in terms of the number of training sessions required to reach criterion: RK required 278 sessions (3336 trials) over 149 days of training (mean 1.86 sessions/day), RB required 265 sessions (3180 trials) over 209 days of training (mean 1.27 sessions/day), EM required 234 sessions (2808 trials) over 175 days of training (mean 1.34 sessions/day), and SH required 487 sessions over 271 days of training (mean 1.79 sessions/day). For SH, we interrupted the training when she gave birth and then restarted the training for 261 sessions (3132) over 143 days (mean 1.81 sessions per day).

Test 1: novel arrangements

Figure 3 shows each subject’s performance at the end of training A (black line), and in test 1 (red line) as percentage of accuracy (number of correct choices/total number of trials \times 100).

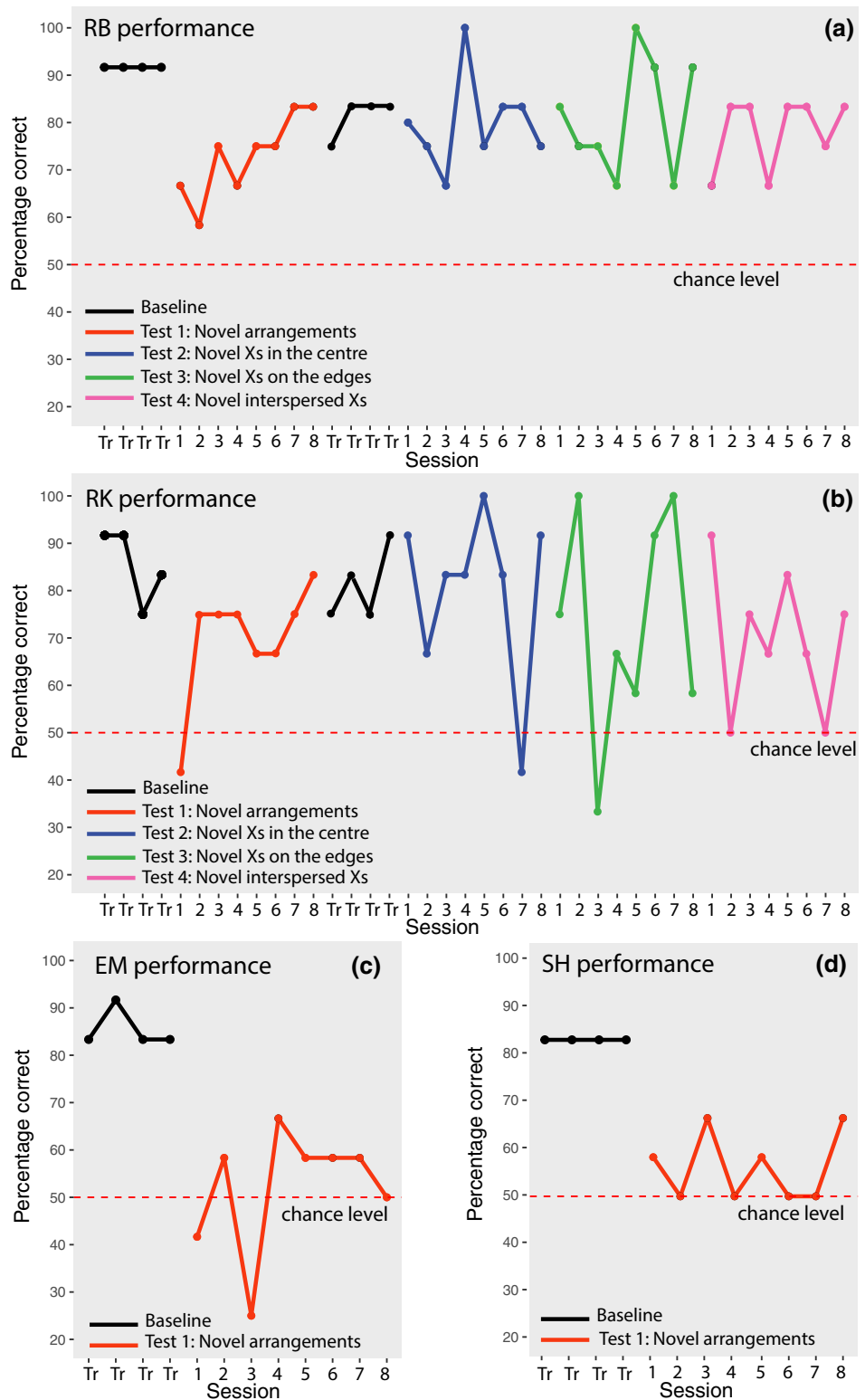
We calculated the number of correct choices for each subject in the first four and eight test sessions (48 and 96 trials, respectively) of test 1 and tested whether the number of correct choices was significantly different from chance with a two-tailed binomial test.

In the first four sessions, two subjects—RK and RB—performed significantly better than chance (RK: 32/48 correct choices, 67%, $p = 0.029$; RB: 32/48 correct choices, 67%, $p = 0.029$), while two subjects—EM and SH—did not (EM: 23/48 correct choices, 48%, $p = 0.885$; SH: 27/48 correct choices, 56%, $p = 0.471$). Similarly, in the first eight sessions, only two subjects performed significantly better than chance: RK: 67/96 correct choices, 70%, $p < 0.001$; RB: 70/96 correct choices, 73%, $p < 0.001$; EM: 50/96 correct choices, 52%, $p = 0.76$; and SH: 54/96 correct choices, 56%, $p = 0.26$.

These results license the conclusion that RK and RB (but not the other two subjects, EM and SH) were able to use the experience gained during training A to successfully distinguish between novel consistent and inconsistent stimuli. Thus, at least two cotton-top tamarins could learn a positional rule as $A(X)^nB$ and generalize this regularity to novel arrangements.

The positive performance of RK and RB is noteworthy considering that, during the training, subjects had the previous experience with only a small set of stimuli ($n = 13$ token combinations), which then increased to a set

Fig. 3 Each panel shows the performance of a subject: **a** for RB **b** for RK **c** for EM **d** for SH. The baseline criterion (black line, Training sessions = Tr) and the first eight sessions of each test are visualized: in test 1 (red line) subjects were probed with novel arrangements of As and Bs, in test 2 (blue line) with novel central Xs, in test 3 (green line) with novel Xs on the edges, in test 4 (pink line) with novel Xs within the array (color figure online)



of 24 novel exemplars presented in test 1, and that before test 1 tamarins had never encountered substitutions with different tokens of the same category. The performance in the generalization test for these subjects, though, was significantly worse than in the baseline, as shown with

a Chi-square test: $\chi^2_{RB} = 9.09, p = 0.002$; $\chi^2_{RK} = 4.63, p = 0.031$). We ran further probe tests to investigate the encoding of the regularity in RB and RK. Since EM and SH did not show signs of generalization with the novel stimuli, and no more novel configurations of As and Bs

were available, we dropped these subjects from subsequent tests.

Training B and test 2: novel Xs in the center

Before moving to test 2, the two subjects that succeeded in test 1 were trained with 36 sessions identical to those presented in test 1 (training B). The success rate over the 36 sessions for each subject was 328/432 correct choices, 76% for RK, and 315/420 correct choices, 75% for RB.

In test 2, we investigated whether RK and RB were able to generalize to new *X* tokens located between *A* and *B*, in the center of the arrays. The results of the first eight sessions are presented as the blue line in Fig. 3. Both monkeys performed significantly above chance (two-tailed binomial tests): RK made 77/96 correct choices, 80%, $p < 0.001$, and RB made 75/94 correct choices (two trials were aborted, because the subject was not responsive), 80%, $p < 0.001$. The same outcomes were observed in the first four sessions: RK made 39/48 correct choices, 81%, $p < 0.001$; RB made 37/48 correct choices, 77%, $p < 0.001$). Hence, the performance of RK and RB was not disrupted by change in *X* tokens in the center of the arrays, suggesting that these monkeys were not using the absolute position of the tokens to make their choices and that the distractor *X* tokens in central position did not affect their choices.

Test 3 and test 4: novel Xs on the edges and between As, Bs, and Xs

In test 3 (Fig. 3, blue line), novel *X*s were added on the edges of the arrays, thus changing the absolute position of *A*s and *B*s with respect to the edges. In the first 8 sessions (96 trials), RK scored 70/96 correct choices, 73%, $p < 0.001$ and RB scored 78/96 correct choices, 81%, $p < 0.001$. Similarly, considering the first four sessions, RK made 33/48 correct choices, 69%, $p = 0.013$, and RB made 36/48 correct choices, 75%, $p < 0.001$). Hence, tamarins responded above chance level even when the absolute position of *A*s and *B*s was changed with respect to the edges.

In test 4 (Fig. 3, green line) at least one *X* was located between *A* or *B* and the central *X*s. In the first eight sessions of test 4, RK made 67/96 correct choices, 70%, $p < 0.001$; RB made 75/96 correct choices, 78%, $p < 0.001$. In the first four sessions, RK made 34/48, 71%, $p = 0.006$; RB made 36/48 correct choices, 75%, $p < 0.001$). Hence, the performance of RB and RK was not disrupted by the insertion of novel *X*s that changed the absolute position of *A*s and *B*s with respect to the center of the array and that made the stimuli asymmetrical. Both monkeys appeared to use a relative position strategy to solve the task. Moreover, these results show that monkeys could process long-distance dependencies, namely, relations between non-adjacent items.

Learning during the test and differences between test stages

We analyzed the performance of RB and RK during test 1 to identify evidence of learning. RB showed a significant increase of performance during the test ($F_{1,6} = 15$, $p = 0.008$), while RK had an abrupt decrease in performance in the first session, but did not show a significant effect of learning ($F_{1,6} = 3.38$, $p = 0.116$), as apparent in Fig. 4.

New stimuli were presented for the first time in the first sessions of each test; for this reason, looking at the performance in the first session is important to understand whether generalization is immediate. Overall, in the first sessions of the tests, monkeys did not perform significantly worse than in subsequent test sessions (Chi-squared test, RB: Chi-squared = 0.112, $p = 0.737$; RK: Chi-squared = 0.017, $p = 0.896$). Moreover, there was no increase in correct responses in subsequent test sessions (Spearman's correlation, RB: $t_{30} = 1.56$, $p = 0.129$; RK: $t_{30} = -0.212$, $p = 0.834$). This evidence suggests that the eight sessions are representative of the overall performance, although the dramatic decrease in performance of RK in the first session of test 1, and the learning curve of RB in test 1 indicates that generalization was not complete.

To test whether the performance of RK and RB differed across test stages, we used a Chi-squared test comparing correct vs incorrect choices in the four tests. RB and RK did not make significantly more incorrect choices across the four different test stages (RB: Chi-squared₃ = 2.344, $p = 0.504$; RK: Chi-squared₃ = 3.457, $p = 0.326$), see Fig. 5.

Analyses of responses to inconsistent stimuli

We ran further analyses to investigate the individual strategies used by the subjects that succeeded in test 1 and went

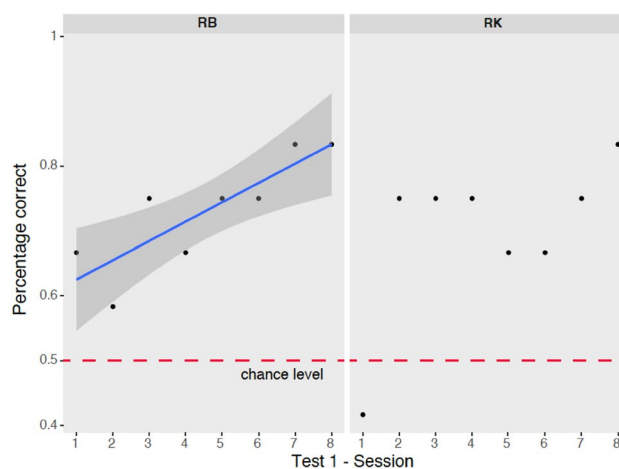


Fig. 4 Performance of RB and RK in the subsequent sessions of test 1

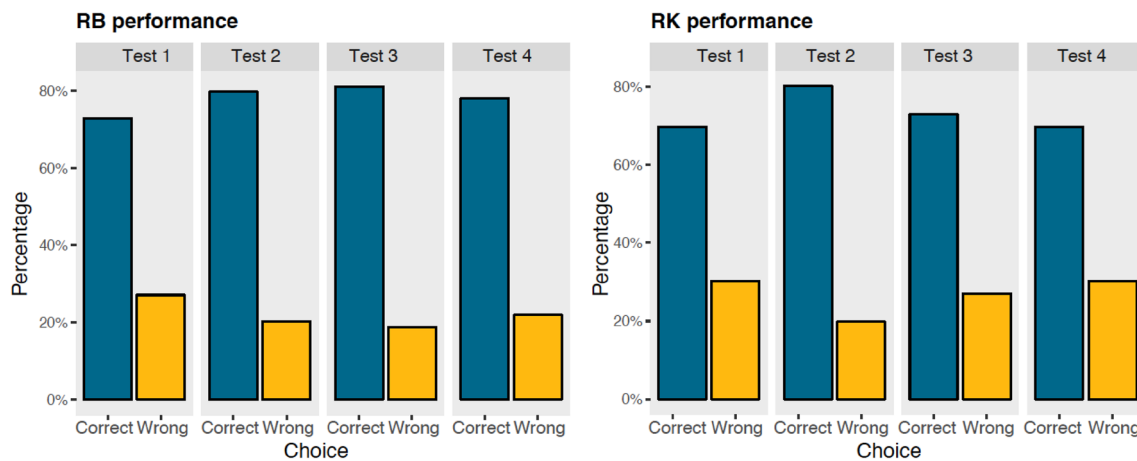


Fig. 5 Overall performance (percentage of correct and wrong choices) by test stage for RB (left panel) and RK (right panel)

through the other test stages, by looking at the pattern of responses to arrangements which were inconsistent with the $A(X)^nB$ grammar in the first eight sessions of each test stage.

To investigate the presence of difficulties or enhanced performance in the presence of specific tokens, we analyzed the responses to each A (Fig. 6) and B token (Fig. 7) presented in ungrammatical stimuli. As shown in Fig. 4, both RB and RK performed significantly worse with inconsistent arrangements that contained A_3 (RB: Chi-squared = 13.737, $p < 0.001$; RK: Chi-squared = 14.462, $p < 0.001$) than with other tokens. During training, we had not used the A_3 token in any unrewarded arrangements (see Table 1) to test for generalization to items presented in novel positions. Hence, the specific failure with A_3 exhibited by both monkeys suggests an incomplete generalization of the positional $A(X)^nB$ rule, focused only on specific items rather than A and B categories. This might be due to the selective inexperience with A_3 as an unrewarded token. This limited generalization suggests that monkeys, similarly to what is observed in pigeons (Herbranson and Shimp 2008), use the strength of string fragments, namely, the resemblance to previously rewarded–unrewarded stimuli, in this task.

We did not find learning difficulties with any other token, but the performance of each monkey was enhanced in the presence of some tokens, although these effects were not consistent across subjects. These are the results with A tokens (Fig. 6) for RB: A_1 Chi-squared = 0.110, $p = 0.740$; A_2 : Chi-squared = 1.490, $p = 0.222$; and for RK: A_1 Chi-squared = 0.110, $p = 0.749$; A_2 Chi-squared = 0, $p = 1$. These are the results for B tokens (Fig. 7) for RB: B_1 Chi-squared = 4.938, $p = 0.026$ with a significant performance enhancement; B_2 Chi-squared = 1.276, $p = 0.259$; B_3 Chi-squared = 3.595, $p = 0.058$ with a trend for performance enhancement; B_4 Chi-squared = 1.292, $p = 0.255$, and for RK: B_1 Chi-squared < 0.001, $p = 1$; B_2 : Chi-squared = 0.0008, $p = 0.977$; B_3 Chi-squared = 0.358,

$p = 0.550$; B_4 Chi-squared = 3.235, $p = 0.071$ with a trend for performance enhancement.

We analyzed the effect of different arrangement types on violations: presence/absence of transpositions (swapped position of A s and B s), presence/absence of identity violations (two identical A or B tokens), presence/absence of violations located on the edges or on the center (Fig. 8).

If monkeys were taking into account the relative position of the tokens, we would have expected a better performance with transpositions, in which the A and B token is swapped, thus producing a double violation, compared to substitutions, in which only one token is located in the incorrect position. On the contrary, we could have observed a pattern similar to the transposed-letter effect observed in human literates (Dunabeitia et al. 2014; Grainger 2008; Perea and Lupker 2004) and pigeons (Scarf et al. 2016), in which ungrammatical letter strings (non-words) obtained by transposing adjacent letters in a grammatical letter string (word) induce misclassifications. Comparing the performance of transpositions vs substitutions, we did not observe any significant enhancement or decrease in performance with transpositions compared to substitutions, although RB had a trend for enhancement with transpositions (RB Chi-squared = 2.773, $p = 0.096$ with a trend for enhancement; RK Chi-squared = 0.042, $p = 0.838$). Hence, differently from pigeons (Scarf et al. 2016), tamarins' performance was not negatively affected by transposition, and potentially slightly enhanced.

We tested whether rejecting arrangements with substitutions with two identical tokens was easier than rejecting other violations of the grammar, and this was not the case for RB (Chi-squared = 1.059, $p = 0.303$), while there was a trend for enhancement in RK (Chi-squared = 2.653, $p = 0.056$). Violations on the edges (RB Chi-squared = 0.552, $p = 0.457$; RK Chi-squared = 0.004, $p = 0.947$) and on the center (RB Chi-squared = 0.097, $p = 0.756$; RK Chi-squared = 0.014,

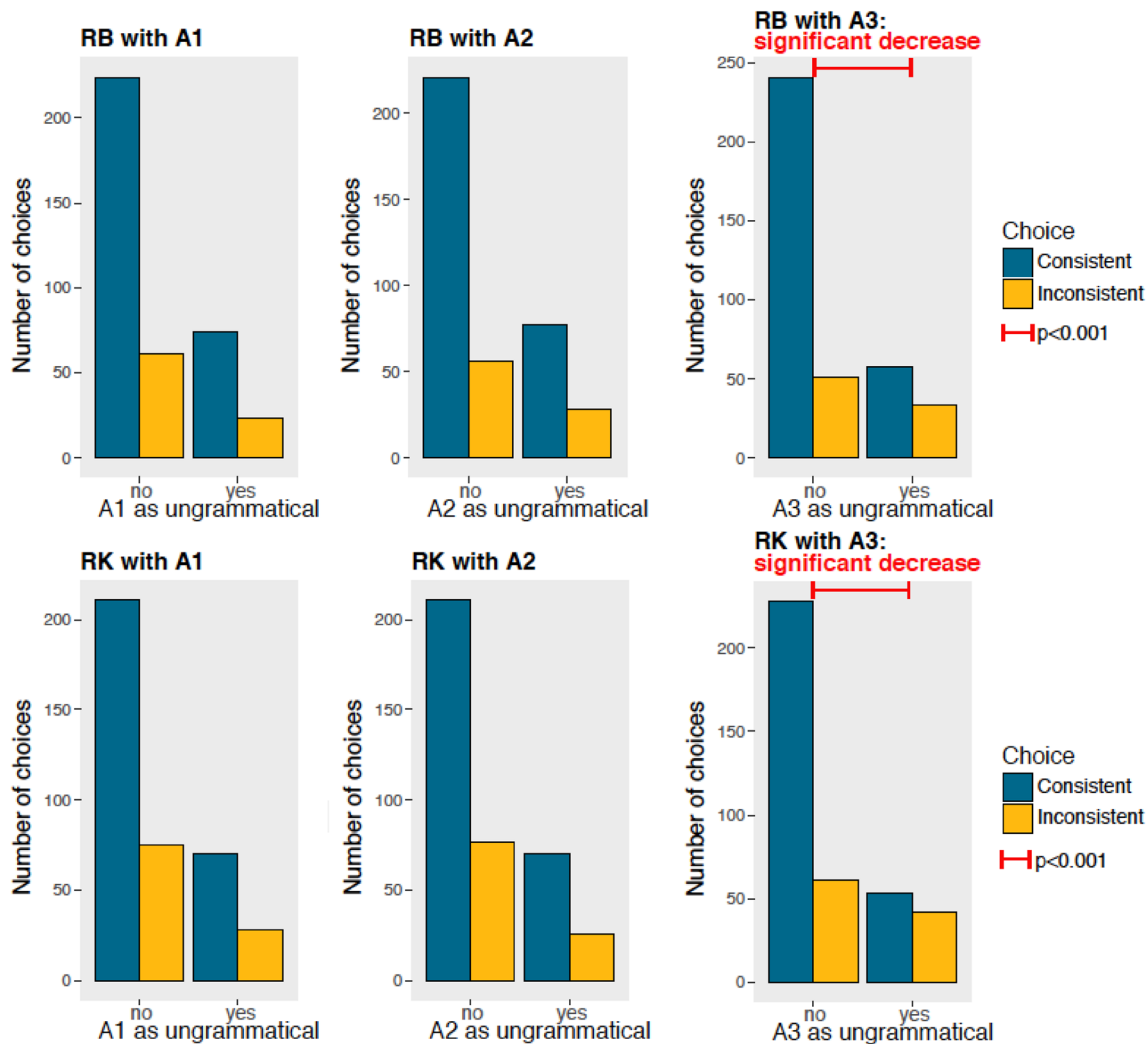


Fig. 6 Each chart shows the performance of a subject (RB on the top panels, RK on the bottom panels) when each A token was present or absent in the ungrammatical arrays. Blue bars indicate correct choices, consistent with the grammar; yellow bars indicate wrong

choices, inconsistent with the grammar. Both RB and RK performances were significantly lower when A₃ tokens were presented in ungrammatical arrangements, suggesting a similar encoding of the regularity (color figure online)

$p = 0.878$) were not significantly different than other violations. Overall, the lack of enhanced performance with any class of violation suggests that monkeys used a strategy based on configurational encoding, rather than an encoding based on the analysis of the position of each token, to respond to novel stimuli.

General discussion

Detecting in which position elements occur relative to one another is important in many domains, such as causal reasoning, language processing, and animal communication (e.g., ten Cate and Okanoya 2012), orthographic processing (Dunabeitia et al. 2014; Scarf et al. 2016; Ziegler et al. 2013), spatial navigation and foraging (Rugani et al. 2010;

Vallortigara and Zanforlin 1986), and action planning (Raby et al. 2007). It is not clear, though, which representations and computational mechanisms different species use to encode positional relationships. Comparative studies have just started to address this issue in primates (Grainger et al. 2012; Sonnweber et al. 2015), rodents (Murphy et al. 2004), and avian species (Chen and ten Cate 2017; Chen et al. 2014; Ravnani et al. 2015; Rosa-Salva et al. 2018; Scarf et al. 2016). While some studies have shown successful mastery of the intended pattern, others suggest that animals sometimes focus on a different level of granularity (Ravnani et al. 2015; Wakita 2019). As a parallel in the auditory domain, most humans have relative pitch perception, while many non-human species in similar conditions exhibit absolute pitch (Honing 2019; Bregman et al. 2016). When it comes

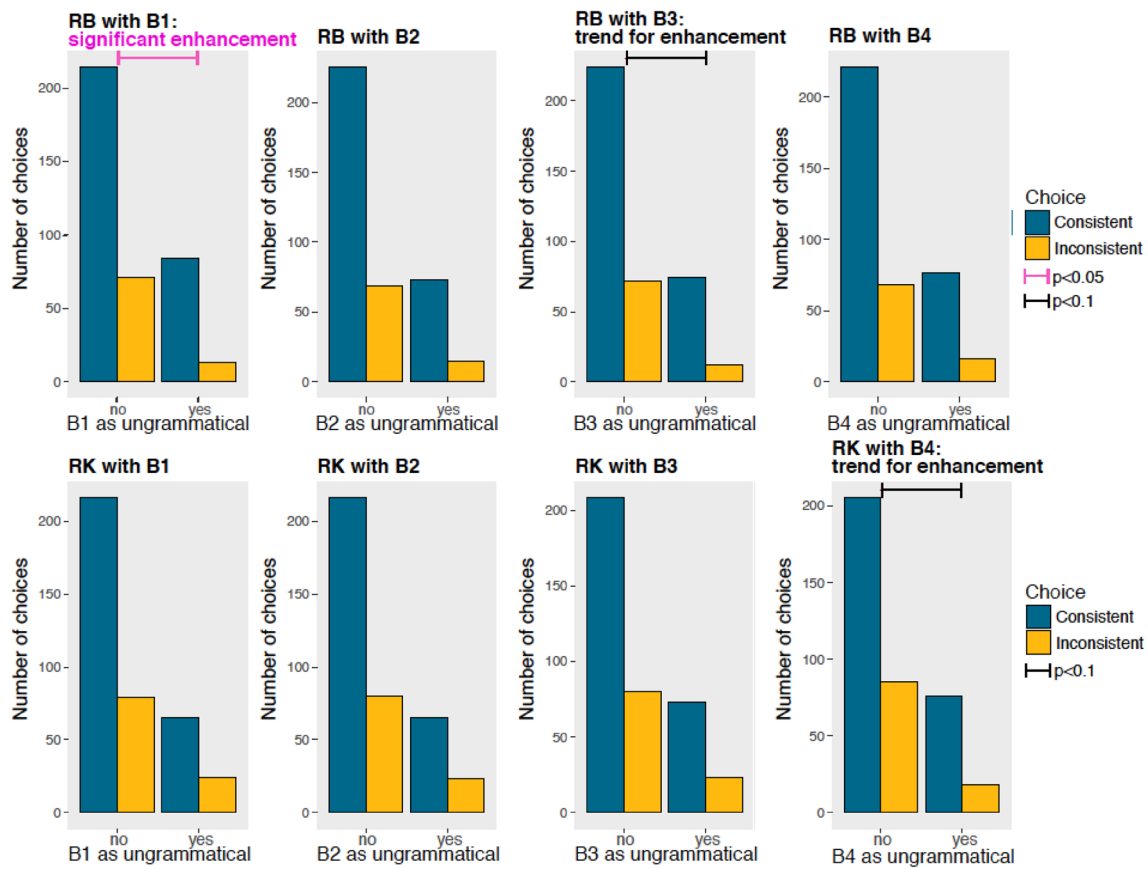


Fig. 7 Each chart shows the performance of a subject (RB on the top panels, RK on the bottom panels) when each *B* token was present or absent in the ungrammatical arrays. Blue bars indicate correct choices, consistent with the grammar; yellow bars indicate wrong

choices, inconsistent with the grammar. Both RB and RK performances were significantly lower when A_3 tokens were presented in ungrammatical arrangements, suggesting a similar encoding of the regularity

to the comparative study of positional regularities, the cross-species picture is still unclear.

In the present work, we investigated how the positional regularity $A(X)^nB$ is processed in the new world monkey species of cotton-top tamarins. Only stimuli in which tokens of the category *A* are located to the left of tokens of the category *B* are consistent with this rule, while tokens other than *As* and *Bs* are not relevant in determining the grammaticality of stimuli. To reduce the attention and memory load required to solve the task and focus on computational capabilities and generalization, we tested monkeys using visual arrays instead of serially presented stimuli that require sustained attention and memory to be processed. Although training methods might not reflect the use of spontaneous abilities, comparative studies can still reveal the presence of differential strategies used in extraction/detection (Gentner et al. 2006; Sonnweber et al. 2015; Spierings and ten Cate 2016) and production (Fitch 2018; Jiang et al. 2018) of regularities from the environment. Jiang et al. (2018) have recently shown that monkeys require thousands of trials to learn supra-regular grammars that are almost immediately

available to preschool children. Although monkeys and children are exposed to different environment, that can influence their learning strategies and skills, strategies and ease in regularity extraction and production might hence constitute a computational divide between humans and non-linguistic species.

Two of our trained monkeys were able to learn a positional regularity between non-adjacent items such as $A(X)^nB$. After being trained on a subset of stimuli, two subjects were able to generalize to novel arrangements the distinction between items consistent and not consistent with the target rule. The limited evidence provided during the training was sufficient for monkeys to tell apart the positional role of *A*, *B* and distractor *X* tokens but with important limitations. The task was solved by half of the sample, and it also took thousands of trials of training to reach the criterion for the test. This could be in part explained by the fact that the bidimensional stimuli with the specific shapes chosen for the test might have little ecological relevance or perceptual appropriateness (Ravignani et al. 2019) for tamarins, but does not completely account for the performance in the generalization

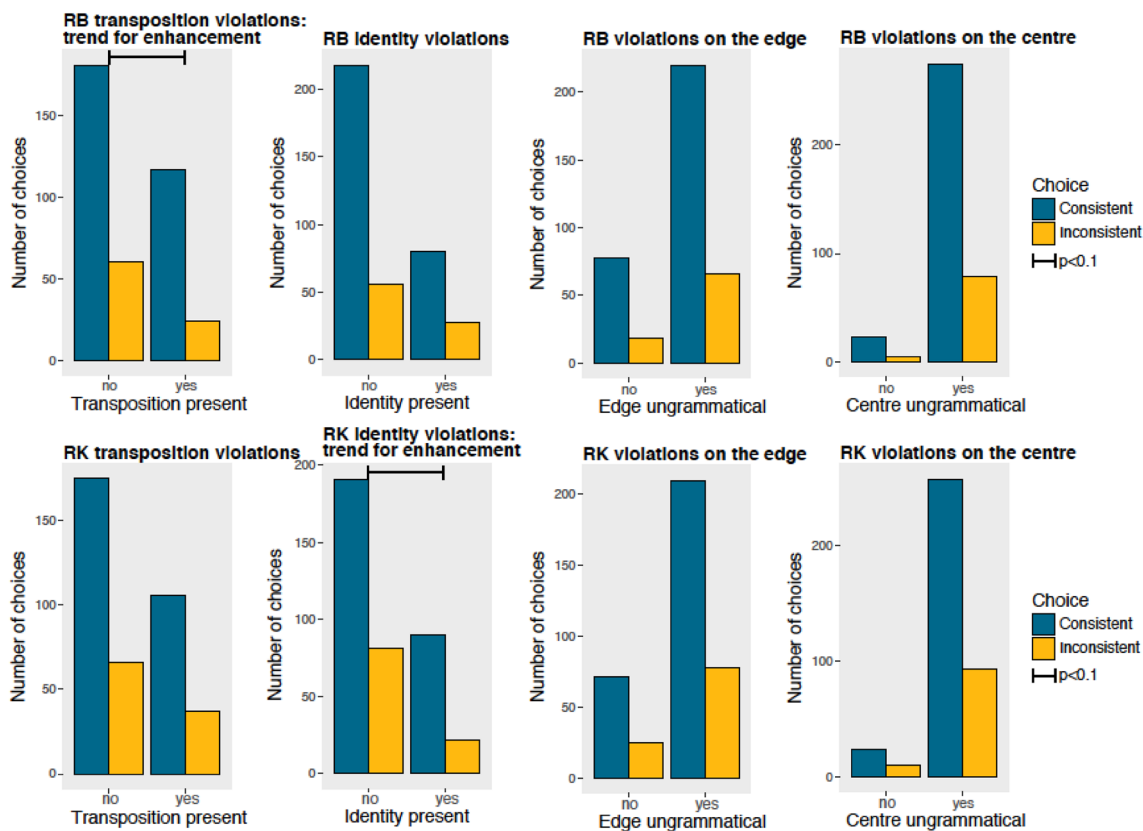


Fig. 8 Each chart shows the performance of a subject presented with different violations: presence/absence of transpositions (swapped position of *A*s and *B*s), presence/absence of identity violations (two

identical *A* or *B* tokens), and presence/absence of violations located on the edges or on the center

tests, after reaching the learning criterion. The performance at test was significantly worse than in the baseline, suggesting a limited generalization, in line with other studies conducted in non-human species (Herbranson and Shimp 2008; Jiang et al. 2018; Spierings and ten Cate 2016).

In another work, Grainger et al. (2012) trained baboons to categorize novel visual arrays of stimuli composed by letters arranged according to a statistical positional regularity. Monkeys were able to acquire adjacent dependencies (relations between adjacent items) over letter bigrams within the arrays and coded the word vs non-word stimuli as sets of letter identities arranged in a specific order. A subsequent experiment with baboons (Ziegler et al. 2013) showed that monkeys, similar to human readers, exhibits the transposed-letter effect (see Grainger 2008), so that their performance was lowered by letter transpositions. In this work, though, it was not clear whether monkeys relied on an absolute or on a relative position strategy. A significant decrease in performance with transpositions mediated by relative position encoding has been noticed in human literates (Dunabeitia et al. 2014; Grainger 2008) and in pigeons trained to orthographical discriminations (Scarf et al. 2016). In our experiments, the absolute position of a token within an array, such

as “ A_1 must be located as first token on the left part of the array”, is defined independently from the identity of other tokens. On the contrary, its relative position, such as “ A_1 must be located to the left of $B_1, B_2, B_3,$ or B_4 ”, depends on the specific identity and position of other tokens. These two alternative strategies of encoding can be probed changing the absolute position of *A*s and *B*s with respect to the edges and the center of the array by inserting novel *X*s in different positions within the visual arrays. If during the training tamarins had encoded the absolute and not the relative position of *A*s and *B*s, they were expected to fail when the absolute position of *A*s and *B*s was changed. On the contrary, tamarins’ performance was not disrupted when novel *X*s were added in the center of the arrays and when the absolute position of *A*s and *B*s was changed with respect to the edges and the center of the array. We can hence conclude that tamarins did not rely on the mere absolute position of *A*s and *B*s. The fact that tamarins’ performance was not compromised by transpositions suggests that they might have used a visual similarity strategy to solve the task. This conclusion is supported by the fact that both successful subjects had a significantly lower performance with ungrammatical stimuli

that contained a token— A_3 —never presented in unrewarded stimuli during the training.

Overall, we documented a preferential encoding of the relative position in successful subjects. Preferential encoding of relative vs absolute position has been observed for instance in baboons (Ziegler et al. 2013) and pigeons (Scarf et al. 2016). Relative rather than absolute encoding in spatial positions had been previously shown also in chicks of the domestic fowl during foraging (Vallortigara and Zanforlin 1986). In a series of experiments, chicks were trained to discriminate between two boxes according to either their relative position to each other or their absolute position (the position with respect to the geometry of the cage or other features of the environment). When the boxes were located close to each other, learning on the basis of the relative position was faster than learning on the basis of absolute position. The advantage of relative position was reduced only when the boxes were located further apart. Similarly, our results obtained with closely located tokens show a preference for relative encoding of visual stimuli presented in simultaneous configurations. It seems that encoding of relative rather than absolute position is a frequent strategy observed across different species and taxa (possibly the default strategy for the encoding of closely located items), irrespective of the possess of language. This study expands our knowledge of preferential encoding strategies of regularities in primates, pointing at some potential limitations of generalization in monkeys. Successful subjects, in fact, used the presence of a specific item that was previously associated only to rewarded trails, and not its position, as a cue for solving the task. This points to the use of a local strategy, rather than to a complete generalization at the category level. Interestingly, this works shows also that new world monkeys are able to process non-adjacent relationships (relations between non-adjacent items), a feature that appears shared across primates (see, for instance, Milne et al. 2016; Sonnweber et al. 2015). Further studies should explore the differences between sensory modalities and clarify the neurobiological and evolutionary basis of strategies used for positional encoding.

Acknowledgements All research was approved by the Animal Care and Use committee at Harvard University. Funds for this research were provided to E.V. by the Harvard Mind Brain and Behavior (MBB) Graduate Student Award. We thank the members of the Cognitive Evolution Laboratory (Harvard University) for help in data collection and comments on the data.

Author contributions EV conceived and designed the experiments, prepared the setup, conducted the experiments, analyzed the data, drafted the manuscript; JRR: prepared the setup, conducted the experiments and revised the manuscript; NSM: prepared the setup, conducted the experiments and revised the manuscript; AR discussed the data and revised the manuscript. All authors gave final approval for publication.

Funding Andrea Ravignani has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant agreement no. 665501 with the research Foundation Flanders (FWO) (Pegasus² Marie Curie fellowship 12N5517N), and a visiting fellowship in Language Evolution from the Max Planck Society.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Bregman MR, Patel AD, Gentner TQ (2016) Songbirds use spectral shape, not pitch, for sound pattern recognition. *Proc Natl Acad Sci* 113(6):1666–1671. <https://doi.org/10.1073/pnas.1515380113>
- Chen J, ten Cate C (2017) Bridging the gap: learning of acoustic non-adjacent dependencies by a songbird. *J Exp Psychol Anim Learn Cogn* 43(3):295–302
- Chen J, van Rossum D, ten Cate C (2014) Artificial grammar learning in zebra finches and human adults: XYX versus XXY. *Anim Cogn* 18:151–164. <https://doi.org/10.1007/s10071-014-0786-4>
- Chomsky N (1956) Three models for the description of language. *IEEE Trans Inf Theory* 2(3):113–124. <https://doi.org/10.1109/TIT.1956.1056813>
- Dunabeitia JA, Orihuela K, Carreiras M (2014) Orthographic coding in illiterates and literates. *Psychol Sci* 25(6):1275–1280. <https://doi.org/10.1177/0956797614531026>
- Endress AD, Carden S, Versace E, Hauser MD (2010) The apes' edge: positional learning in chimpanzees and humans. *Anim Cogn* 13(3):483–495. <https://doi.org/10.1007/s10071-009-0299-8>
- Fiser J, Aslin RN (2001) Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychol Sci* 12(6):499–504
- Fiser J, Aslin RN (2002) Statistical learning of higher-order temporal structure from visual shape sequences. *J Exp Psychol Learn Mem Cogn* 28(3):458–467. <https://doi.org/10.1037/0278-7393.28.3.458>
- Fiser J, Aslin RN (2005) Encoding multielement scenes: statistical learning of visual feature hierarchies. *J Exp Psychol Gen* 134(4):521–537. <https://doi.org/10.1037/0096-3445.134.4.521>
- Fitch WT (2014) Toward a computational framework for cognitive biology: unifying approaches from cognitive neuroscience and comparative cognition. *Phys Life Rev* 11(3):329–364. <https://doi.org/10.1016/j.plrev.2014.04.005>
- Fitch WT (2017) Empirical approaches to the study of language evolution. *Psychon Bull Rev*. <https://doi.org/10.3758/s13423-017-1236-5>
- Fitch WT (2018) Bio-linguistics: monkeys break through the syntax barrier. *Curr Biol* 28(12):R695–R697. <https://doi.org/10.1016/j.cub.2018.04.087>
- Fitch WT, Friederici AD (2012) Artificial grammar learning meets formal language theory: an overview. *Philos Trans R Soc Lond Ser B Biol Sci* 367(1598):1933–1955. <https://doi.org/10.1098/rstb.2012.0103>
- Fitch WT, Friederici AD, Hagoort P (2012) Pattern perception and computational complexity: introduction to the special issue. *Philos Trans R Soc Lond Ser B Biol Sci* 367(1598):1925–1932. <https://doi.org/10.1098/rstb.2012.0099>

- Frank MC, Gibson E (2011) Overcoming memory limitations in rule learning. *Lang Learn Dev* 7(2):130–148. <https://doi.org/10.1080/15475441.2010.512522>
- Gentner TQ, Fenn KM, Margoliash D, Nusbaum HC (2006) Recursive syntactic pattern learning by songbirds. *Nature* 440(7088):1204–1207. <https://doi.org/10.1038/nature04675>
- Ghirlanda S, Lind J, Enquist M (2017) Memory for stimulus sequences: a divide between humans and other animals? *R Soc Open Sci* 4(6):161011. <https://doi.org/10.1098/rsos.161011>
- Grainger J (2008) Cracking the orthographic code: an introduction. *Lang Cogn Process* 23(1):1–35. <https://doi.org/10.1080/0169096701578013>
- Grainger J, Dufau S, Montant M, Ziegler JC, Fagot J (2012) Orthographic processing in baboons (*Papio papio*). *Science* 336(6078):245–248. <https://doi.org/10.1126/science.1218152>
- Herbranson WT, Shimp CP (2008) Artificial grammar learning in pigeons. *Learn Behav* 36(2):116–137. <https://doi.org/10.3758/LB.36.2.116>
- Honing H (2019) The evolving animal orchestra: in search of what makes us musical. MIT Press, Cambridge
- Jiang X, Long T, Cao W, Li J, Dehaene S, Wang L (2018) Production of supra-regular spatial sequences by macaque monkeys. *Curr Biol*. <https://doi.org/10.1016/j.cub.2018.04.047>
- Milne Alice E, Mueller JL, Mannel C, Attaheri A, Friederici AD, Petkov CI (2016) Evolutionary origins of non-adjacent sequence processing in primate brain potentials. *Sci Rep* 6:36259. <https://doi.org/10.1038/srep36259>
- Milne AE, Wilson B, Christiansen MH (2018) Structured sequence learning across sensory modalities in humans and nonhuman primates. *Curr Opin Behav Sci* 21:39–48. <https://doi.org/10.1016/j.cobeha.2017.11.016>
- Murphy RA, Mondragón E, Murphy VA, Fouquet N (2004) Serial order of conditional stimuli as a discriminative cue for Pavlovian conditioning. *Behav Proc* 67:303–311. <https://doi.org/10.1016/j.beproc.2004.05.003>
- Murphy RA, Mondragón E, Murphy VA (2008) Rule learning by rats. *Science* 319(5871):1849–1851. <https://doi.org/10.1126/science.1151564>
- Perea M, Lupker SJ (2004) Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *J Mem Lang* 51(2):231–246. <https://doi.org/10.1016/j.jml.2004.05.005>
- Raby CR, Alexis DM, Clayton NS (2007) Planning for the future by western scrub-jays. *Nature* 445:919–921. <https://doi.org/10.1038/nature05575>
- Ravignani A, Sonnweber R (2017) Chimpanzees process structural isomorphisms across sensory modalities. *Cognition* 161:74–79. <https://doi.org/10.1016/j.cognition.2017.01.005>
- Ravignani A, Westphal-Fitch G, Aust U, Schlumpp MM, Fitch WT (2015) More than one way to see it: individual heuristics in avian visual computation. *Cognition* 143:13–24. <https://doi.org/10.1016/j.cognition.2015.05.021>
- Ravignani A, Filippi P, Tecumseh Fitch W (2019) Perceptual tuning influences rule generalization: testing humans with monkey-tailored stimuli. *I-Perception*. <https://doi.org/10.1177/2041669519846135>
- Reber AS, Lewis S (1977) Implicit learning: an analysis of the form and structure of a body of tacit knowledge. *Cognition* 5(4):333–361. [https://doi.org/10.1016/0010-0277\(77\)90020-8](https://doi.org/10.1016/0010-0277(77)90020-8)
- Rey A, Perruchet P, Fagot J (2012) Centre-embedded structures are a by-product of associative learning and working memory constraints: evidence from baboons (*Papio Papio*). *Cognition* 123(1):180–184. <https://doi.org/10.1016/j.cognition.2011.12.005>
- Rosa-Salva O, Fiser J, Versace E, Dolci C, Chehaimi S, Santolin C, Vallortigara G (2018) Spontaneous learning of visual structures in domestic chicks. *Animals* 8(8):135. <https://doi.org/10.3390/ani8080135>
- Rugani R, Kelly DM, Szelest I, Regolin L, Vallortigara G (2010) Is it only humans that count from left to right? *Biol Lett* 6(3):290–292. <https://doi.org/10.1098/rsbl.2009.0960>
- Scarf D, Boy K, Uber Reinert A, Devine J, Güntürkün O, Colombo M (2016) Orthographic processing in pigeons (*Columba livia*). *Proc Natl Acad Sci* 113(40):11272–11276. <https://doi.org/10.1073/pnas.1607870113>
- Smith KH (1966) Grammatical intrusions in the recall of structured letter pairs: mediated transfer or position learning? *J Exp Psychol* 72(4):580–588. <https://doi.org/10.1037/h0023768>
- Sonnweber R, Ravignani A, Fitch WT (2015) Non-adjacent visual dependency learning in chimpanzees. *Anim Cogn* 18(3):733–745. <https://doi.org/10.1007/s10071-015-0840-x>
- Spierings MJ, ten Cate C (2016) Budgerigars and zebra finches differ in how they generalize in an artificial grammar learning experiment. *Proc Natl Acad Sci*. <https://doi.org/10.1073/pnas.1600483113>
- Stobbe N, Westphal-Fitch G, Aust U, Fitch WT (2012) Visual artificial grammar learning: comparative research on humans, kea (*Nestor notabilis*) and pigeons (*Columba livia*). *Philos Trans R Soc Lond Ser B Biol Sci* 367(1598):1995–2006. <https://doi.org/10.1098/rstb.2012.0096>
- ten Cate C, Okanoya K (2012) Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philos Trans R Soc Lond Ser B Biol Sci* 367(1598):1984–1994. <https://doi.org/10.1098/rstb.2012.0055>
- Vallortigara G, Zanforlin M (1986) Position learning in chicks. *Behav Proc* 12:23–32
- Versace E, Regolin L, Vallortigara G (2006) Emergence of grammar as revealed by visual imprinting in newly-hatched chicks. In: The evolution of language. Proceedings of the 6th International Conference, Rome, 12–15 April 2006
- Versace E, Spierings MJ, Caffini M, ten Cate C, Vallortigara G (2017) Spontaneous generalization of abstract multimodal patterns in young domestic chicks. *Anim Cogn* 20(3):521–529. <https://doi.org/10.1007/s10071-017-1079-5>
- Wakita M (2019) Auditory sequence perception in common marmosets (*Callithrix jacchus*). *Behav Proc* 162:55–63. <https://doi.org/10.1016/j.beproc.2019.01.014>
- Ziegler JC, Hannagan T, Dufau S, Montant M, Fagot J, Grainger J (2013) Transposed-letter effects reveal orthographic processing in baboons. *Psychol Sci* 24:1609–1611. <https://doi.org/10.1177/0956797612474322>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.