Psychological Research
DOI 10.1007/s00426-011-0328-5

ORIGINAL ARTICLE

Artificial grammar learning in children: abstraction of rules or sensitivity to perceptual features?

Arnaud Witt · Annie Vinter

Received: 30 December 2010 / Accepted: 24 February 2011

Abstract We examined sensitivity to grammatical sequences of colors in an artificial grammar learning task in a sample of 120 children aged between 5 and 8 years. The aim of the experiment was to test whether the children would preferentially learn the specific salient features of the items they were exposed to or the rules that generated these items. The children were divided into two experimental groups (identical grammar but training items differing in their surface features) and a control group (random items). The results showed that regardless of age, participants learned the most frequent salient features of the items, as well as some kind of abstract relational information. However, the 8-year-olds presented a more complex result profile, with one of the experimental groups apparently developing sensitivity to grammatical rules. These results are discussed with reference to the main current models of implicit learning. Overall, the results provided more support for stimulus-specific processing models than for rule-based models.

Introduction

The renewed interest shown by cognitive psychologists in learning in recent decades has led to a growing body of research devoted to the study of implicit learning (e.g., Reber, 1993; Stadler & Frensch, 1998; Jiménez, 2003). Implicit learning (IL) covers all forms of unintentional learning in which, as a consequence of repeated experience, an individual's behavior becomes sensitive to the structural features of an experienced situation, even though the individual is not told at any time to learn anything about this situation, and without the adaptation being due to an intentional exploitation of certain items of explicit knowledge (Perruchet & Vinter, 1998). A prototypical paradigm for the study of IL is the artificial grammar learning (AGL) paradigm (Reber, 1967). In this paradigm, participants are exposed to a subset of grammatical strings generated by a finite state grammar which defines transition rules between elements. Participants know nothing about the existence of the grammar. They are then tested to see whether they can nevertheless discriminate between new grammatical strings and non-grammatical strings. Results convincingly show that participants recognize grammatical strings at a level significantly above chance as if they had discovered the rules of the grammar, although they are usually unable to make any statements about these rules (see Pothos, 2007, for a review). Most of the studies in this area of research are laboratory studies involving adult participants. This fact, however, should not mask the extent to which IL is involved in many children's acquisitions. Implicit learning has been considered to be responsible for at least some aspects of first- and second-language learning (Carr & Curran, 1994; Chandler, 1993), reading and spelling acquisitions (Pacton et al., 2001), or adaptation to the physical constraints of the world (Krist et al., 1993). Nevertheless, the current literature on IL tells us little about the type of knowledge children acquire about their environment as a result of these processes. The aim of the present study was to test the predictions made by different models of IL in order to characterize the contribution of IL processes to the construction of knowledge in children.

Various theoretical models of IL are described in the literature. We will focus on three of them and illustrate them in the light of the AGL paradigm. Reber (1967, 1989)

A. Witt · A. Vinter (✉)
University of Bourgogne, LEAD-CNRS 5022, Pôle 2AFE,
21000 Dijon, France
e-mail: annie.vinter@u-bourgogne.fr

has claimed that IL results from a highly sophisticated unconscious processor which is devoted to the abstraction of rules. The acquired knowledge refers to the grammatical structure of the material itself, which is strictly independent of statistical information contained in the material seen during training. By contrast, Brooks (1978) and Vokey and Brooks (1992) have proposed the exemplarist theory which suggests that participants memorize whole items or parts of them during training. The stored exemplars would supply sufficient knowledge with which the items proposed during the test phase are compared. This conception has been extended by the similarity assumption (Brooks & Vokey, 1991) which holds that grammatical judgments are based on the degree of similarity between the stored items and those presented during testing. A more recent model, the Self-Organizing Consciousness (SOC) model, was developed by Perruchet and Vinter (2002) and incorporates the main principles of the fragmentarist view (Perruchet, 1994). According to this view, participants automatically engage attentional processes guided by their knowledge of the situation and the properties of the seen items, such as their statistical properties or salient perceptual features. The more salient a feature is, the more likely it is to attract attention and consequently create a cognitive unit. The more frequently this feature occurs, the more strongly the cognitive unit will be consolidated. Thus, IL processes directly shape participants' behavior and their perceptual understanding of the situation.

These models diverge on several points, including the nature of the acquired knowledge. In the abstractionist view, this is linked to the grammatical deep structure whereas, in the other conceptions, it is associated with the properties of the stimuli. Manza and Reber (1997) thought of this opposition as contrasting a position where "the representation is intimately tied to the physical form of the stimulus" with a position where the representations "are devoid of specific surface information but retain some underlying structural knowledge of stimuli" (p. 98). According to Reber (1993), abstraction processes operate on the legal co-occurrences seen in the training material regardless of their frequency, which relates to the superficial properties of the material. This proposition offers an interesting way to differentiate between rule-based and stimulus-specific processing models of implicit learning. In the present experiment, from the same artificial grammar two different sets of items were generated, that contained the same grammatical co-occurrences at the bigram level but differed in terms of stimulus-specific features, namely the frequencies of salient grammatical units. Frequency-independent hypotheses can be derived from rule-based models, by contrast with stimulus-specific processing models. Indeed, as claimed by Smith, Langston and Nisbett (1992, p. 70) in their essay to define the criteria for abstract rules use, "the more familiar an item is, the less likely it is to be encoded abstractly. This is because familiarity often rests on frequency, and frequent presentations of an item might lead one to represent it in terms of its specific content".

The salient surface features focused on here relate to the presence of adjacent (ADJ) and non-adjacent (non-ADJ) repetitions in the items. Children and adults, as well as animals, are known to be highly sensitive to repetitions of events in their environment (e.g., Endress, Dehaene-Lambertz, & Mehler, 2007; Monaghan & Rowson, 2008; Gallistel, 1990, 2000). The two series of grammatical items generated from an artificial finite state grammar inspired from Reber (1967) contained a large number of ADJ repetitions (2 ADJ rules presented 4 times) and a few non-ADJ repetitions (1 non-ADJ rule presented once) for the ADJ series, and a large number of non-ADJ repetitions (3 non-ADJ rules, presented 4, 3 and 1 times respectively) and a few ADJ repetitions (2 ADJ rules presented once) for the non-ADJ series. We also followed Perruchet and Reber's (2003) recommendation to add a control group that was trained in conditions as close as possible to the training conditions used for the experimental groups. The control group was exposed to random items (RAND series), which contained ADJ or non-ADJ units as determined by chance. We implemented the grammar with colors and generated sequences of colors to make flags. The current literature on the automatic encoding of colors suggests that this type of learning process operates efficiently given that attention is directed towards colors (Patel et al., 1999). When the color information directly referred to the objects with which the children interacted, they correctly recalled the color of the objects (Ling & Blades, 1996, 2002). Thus, in our experiment, children were randomly assigned to a group trained with the ADJ series, the non-ADJ series or the RAND series. What predictions concerning IL effects in these groups can be made on the basis of the theoretical models presented above?

A strong rule-based position (Hypothesis 1) claims that the knowledge acquired during training is independent from the specific features of the material presented in the acquisition phase. One crucial argument has been provided by Reber and Lewis (1977) who reported that participants' performance at test correlated with the whole set of grammatical items that the grammar made it possible to generate and not with the specific subset of items seen during training. In our experiment, ADJ and non-ADJ rules were present in both series, but occurred at different frequencies. As a genuine rule-based perspective makes clear, sensitivity to rules cannot be the result of information frequency. The simple presence of the two types of units in each grammatical series should be sufficient for rule abstraction. Thus, according to this theory, no differences were expected between the ADJ and the non-ADJ groups and the theory only predicts differences between each experimental group and the control group (RAND group). Both experimental groups should produce ADJ as well as non-ADJ units, at a significantly higher level than the control group. However, a mild rule-based view (Hypothesis 2) as developed by Reber (1993) could argue that participants acquired surface-independent knowledge in the form of frequency-independent rules, such XX (ADJ) or XaX (non-ADJ), on the basis of the seen material, not on the basis of a hypothetical grammar that exists only from the experimenter's point of view. In the case of the non-ADJ training condition for instance, participants would form 3 rules for allowed non-ADJ repetitions and 2 rules for ADJ repetitions. The ADJ training condition would lead to the formation of 2 rules for allowed ADJ repetitions and 1 rule for non-ADJ repetition. A "random walk in that person's world landscape"[1] would generate a larger production of ADJ than of non-ADJ repetitions in the ADJ training condition (ratio 2/1) and slightly more non-ADJ than ADJ repetitions in the non-ADJ training condition (ratio 3/2). These productions should still be significantly higher than those observed in the control group. By contrast, stimulus-specific processing models (Hypothesis 3) predict much greater differences in learning effects between the ADJ and non-ADJ groups, reflecting the differences in frequency information (ADJ group: ratio 8/1 and non-ADJ group: ratio 8/2), and only partial differences between the experimental groups and the RAND group. If participants memorize exemplars (Brooks, 1978), it is very likely that these exemplars embed adjacent repetitions in the ADJ group and non-adjacent repetitions in the non-ADJ group. If saliency guides the participants' attentional focus then the elements involved in the repetitions should become easily associated and reinforced because of their frequency (Perruchet & Vinter, 2002). Participants should therefore perceive the material as a function of the content of the salient units they have been exposed to. The ADJ group should produce more ADJ units than the non-ADJ group and control group, but no more non-ADJ units than the control group. The non-ADJ group should produce more non-ADJ units than the ADJ group and control group, but no more ADJ units than the control group. Figure 1 summarizes the three hypotheses that can be drawn from the current literature. We used a generation test in our experiment in order to gain a more direct understanding of the knowledge which participants had as a result of their implicit learning experience.

If rule-based models contrast with stimulus-specific learning models with regard to the previous predictions, it is worth pointing out that these models lead to identical predictions with regard to the question of knowing whether the very structure of repetition (XX or XaX) can be learned independently of color encoding (Yellow-Yellow or Yellow-green-Yellow for instance). The three models described above consider this possibility. The rule-based abstractionist model for instance takes transfer of the underlying rules as evidence for participants having learned the grammar (e.g., Gomez & Gerken, 1999). The similarity between AXX and BYY, for instance, is certain (Brooks & Vokey, 1991). When participants are confronted with units like BYY, for instance, the experience resulting from the attentional processing of such units is probably not limited to the sum of its elements, but instead integrates some direct perception of the overall structure (Perruchet & Vinter, 2002). It is therefore very likely that the repetitive structure will be encoded directly. This shows that each of these models envisages some forms of abstraction in the course of an implicit learning episode.

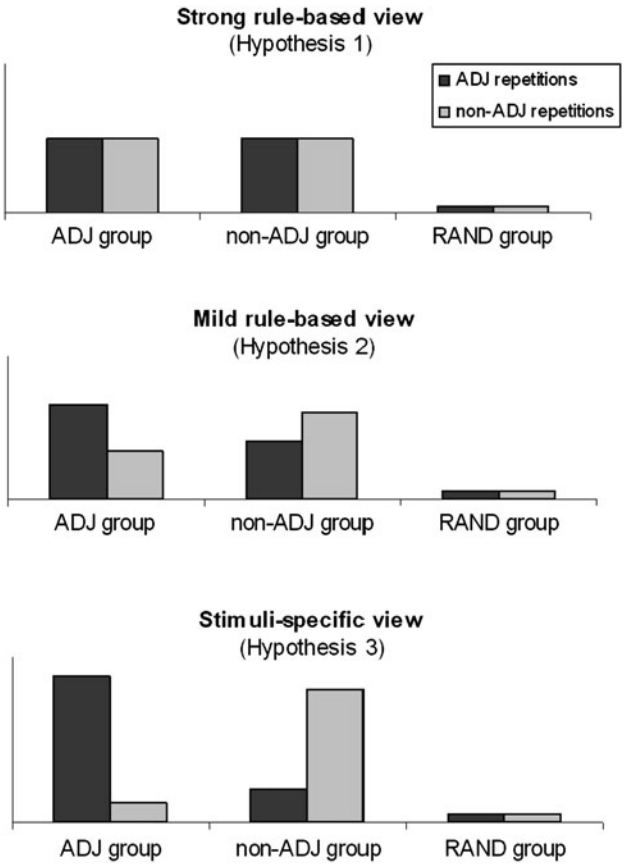


Fig. 1 Illustration of the three hypotheses drawn from the main models of implicit learning (Hypothesis 1: Strong rule-based view; Hypothesis 2: Mild rule-based view; Hypothesis 3: Stimuli-specific view)

These predictions were tested in an experiment carried out with children aged between 5 and 8 years and which involved a generation test (building colored flags) performed in response to implicit instructions. What was the

---

[1] We sincerely thank an anonymous reviewer for this nice metaphor and argument.

main interest of testing children in this context? Including different age groups was an appropriate way of testing whether the same learning processes applied at the various ages, or whether rule-based models and item specific learning models account from implicit learning performance, depending on the participant's age. Indeed, it could be argued that older children are more likely than younger ones to engage in rule-based processes. Their performance should therefore be closer to the predictions resulting from the rule-based positions than that of younger children. The age period between 5 and 8 years was selected for several reasons. First, at these ages children probably possess minimal knowledge of the flags that represent the different countries around the world, limiting that this kind of knowledge affected their production, since the implicit instructions delivered in test did not constrain the generation task. Secondly, we selected the largest developmental period where our experimental context was still attractive to children. Finally, during early childhood, children are exposed to a relatively high level of repetition (like in nursery rhymes for instance) and we were interesting in investigating how much repetition may influence their cognitive system.

Method

Participants

One hundred and twenty Caucasian kindergarten pupils and first and second graders (61 female and 59 male), of between 5 and 8 years of age, participated in the experiment. Four age groups corresponding one-to-one to each of the four school grades were constructed ($N = 30$ per age group). Within each age group, the participants were randomly subdivided into three groups: two experimental groups (ADJ group, $N = 40$, and non-ADJ group, $N = 40$) and a control group (RAND group, $N = 40$). None of the children was educationally advanced or retarded or suffered from attentional or intellectual deficits. Their vision was normal or corrected to normal and a rapid verification prior to the experiment revealed that they were able to discriminate and name the five colors used in the experiment. Written parental consent was obtained for each child. The study has been performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. Table 1 provides the characteristics of the groups.

Materials

The material consisted of a computer game involving 3-, 4- and 5-color flags, as illustrated in Fig. 2. Throughout the task, the instructions were given by a prerecorded voice. Three types of series of flags were built: two grammatical series (ADJ and non-ADJ series) and a random series (RAND series). All the series included eight flags: two 3-color flags, three 4-color flags and three 5-color flags. At test, three templates representing flags of 3, 4 or 5 colors respectively were used. These contained 25 colored squares (5 blue, 5 green, 5 red, 5 yellow and 5 turquoise), as illustrated in Fig. 2b.

Construction of the grammatical series

The grammatical series were generated by a grammar that combined five colors (blue, green, red, yellow, and turquoise). An example is shown in Fig. 3. The colors were randomly assigned to the positions in the grammar (blue, for instance, could appear in the first, second, third, fourth or fifth position) so that 10 different outcomes of both the ADJ and non-ADJ series were generated. Each

Table 1 Characteristics of the groups in Experiment 1 (non-ADJ = non-adjacent repetitions, ADJ = adjacent repetitions, RAND = random, F = female, M = male)

Age groups	Mean age (years, months)	Groups	Number	Mean age (years, months)	Sex
5 years	4.11 (range: 4.1–5.1)	Non-ADJ	10	4.10	5F-5M
		ADJ	10	4.11	4F-6M
		RAND	10	4.11	4F-6M
6 years	6 (range: 5.8–6.6)	Non-ADJ	10	6	6F-4M
		ADJ	10	6	4F-6M
		RAND	10	6.1	6F-4M
7 years	7.1 (range: 6.9–7.5)	Non-ADJ	10	6.11	5F-5M
		ADJ	10	7.2	5F-5M
		RAND	10	7.3	6F-4M
8 years	8.2 (range: 7.7–8.9)	Non-ADJ	10	7.11	6F-4M
		ADJ	10	8.4	5F-5M
		RAND	10	8.4	6F-4M

child within each group was exposed to one of these outcomes.

The grammar made it possible to generate non-adjacent repetitions (e.g., BYB), as well as adjacent repetitions (e.g., YY). The non-ADJ series consisted of 3 non-adjacent repetitions presented once, 3 or 4 times respectively, 2 adjacent repetitions each presented once, 8 bigrams without adjacent color repetitions and 9 trigrams without non-adjacent color repetitions (for the sake of illustration, the non-ADJ series corresponding to Fig. 3 was: BYB, RGR, RGRT, BYBT, RGYB, BYYBT, BYGYB, RGGRT). The ADJ series were made up of 2 adjacent repetitions presented 4 times each, 1 non-adjacent repetition presented once, 7 bigrams without adjacent color repetitions and 9 trigrams without non-adjacent color repetitions (the ADJ series corresponding to Fig. 3 was: BYY, RGG, BYYG, BYGR, RGGY, BYYBT, RGGYY, RGGYB). The co-occurrences at the bigram level (with or without repetitions) were identical between the two series except that the non-ADJ series included one more bigram (RT for the series corresponding to Fig. 3). More precisely, the ADJ and non-ADJ series shared 90% of the bigrams and 29% of the trigrams. The total numbers of occurrences of non-ADJ repetitions within the non-ADJ series and of ADJ repetitions within the ADJ series were identical (8), while the numbers of ADJ repetitions in the non-ADJ series (2) and of non-ADJ repetitions in the ADJ series (1) were determined by chance. The probability of obtaining an adjacent repetition with 5 colors was calculated in proportion to the number of flags composing a series (8). We thus obtained 2 adjacent repetitions ((5 repeated bigrams/25 possible bigrams) * 8 = 1.6, rounded to 2). The same method was adopted for the non-adjacent repetitions, while taking account of the fact that the repeated color in a

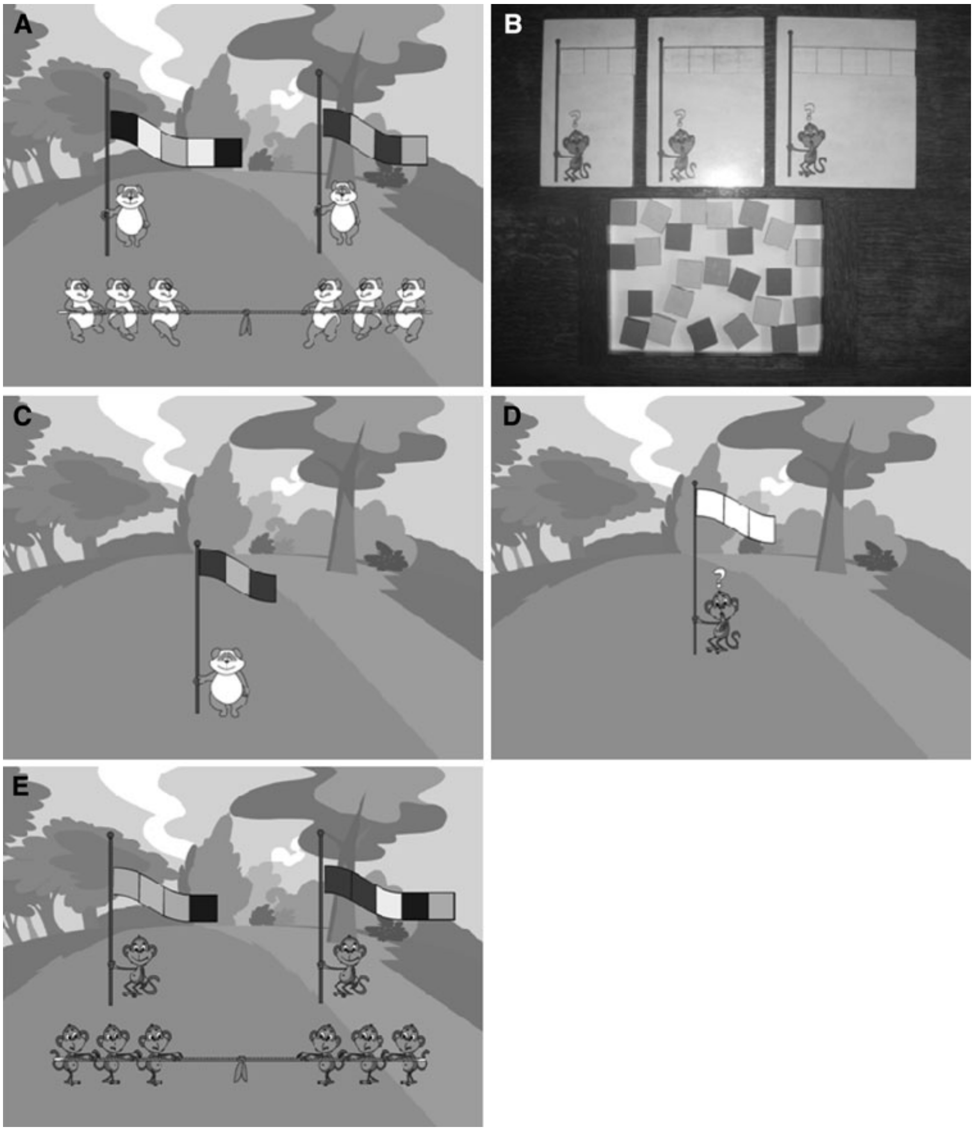
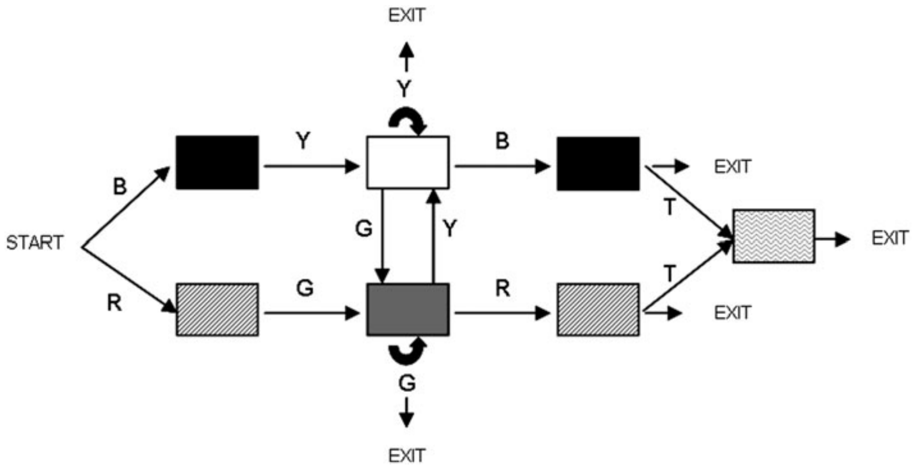


Fig. 2 Illustrations of the video game (**a** tug-of-war tournament; **b** templates and colored squares; **c** presentation of successive colors; **d** implicit generation phase, **e** new flags produced by children)

Fig. 3 An instantiation of the finite state grammar used in the experiment (this grammar was used in both experiments; the position of the colors was variable)



non-ADJ repetition (B_B for instance) could be separated by 1, 2 or 3 elements (B_B, B_ _B, B_ _ _B for instance). This led to the inclusion of 1 non-adjacent repetition in the ADJ series ((20/125 + 80/625 + 400/3215)/3) * 8 = 1.11, rounded to 1). In each series, the 8 grammatical items exposed the children to all the possible paths through the grammar. The statistical coverage of the grammar by the ADJ series, that is the amount of statistical information contained in the items, was .11 and the corresponding value for the non-ADJ series was .30 (see Poletiek & van Schijndel, 2009). These values denote low and high statistical coverages, respectively, following Poletiek and van Schijndel (2009), who showed that rule abstraction should be favored in the case of high coverage condition.

Construction of the random series

Ten outcomes were generated for the random series, with each color occurring in initial position at least once. The RAND series were matched to either the non-ADJ series (half of the series) or to the ADJ series (the other half) in terms of color frequencies. Thus, the RAND series consisted of 2 adjacent repetitions presented once, 1 non-adjacent repetition presented once, 13 or 12 bigrams, and 11 or 13 trigrams (without repetitions), depending on whether they were paired with the ADJ or non-ADJ series respectively (for example, YBG, RGY, GYBR, BGGY, RGBY, TRYYG, YTGBR, GYBYR).

Procedure

Presentation and training

The experimental session comprised a 20 min phase of exposure to the material, followed by a 10 min test phase. The children were comfortably seated in front of a computer and were told that they were going to play a video game. When they were ready, the game started and the children followed the prerecorded instructions: “Hello, today the pandas have organized a “tug of war” tournament. Each team of pandas will show you its pretty flag. Press “start” to see the first team’s flag”. The colors of the flag appeared one at a time, for 500 ms, from left to right. Then, after an interval of 1 s, the children heard the instruction “Now, press “start” to see another team’s flag”. The 8 flags comprising the series appeared successively one at a time (Fig. 2c) in random order. The instructions then continued: “Now, the tournament is going to start. Press “start” to see the first team’s flag” (a flag was displayed, one color at a time, followed by the sound of a trumpet). “Press “start” to see the second team’s flag” (the second flag was displayed). “Now, press “start” to start the match”. The first team of pandas faced the second one (see Fig. 2a). The two flags remained visible until one of the teams won. This procedure was repeated throughout the 16 matches of the tournament. Thus, the children were exposed to the 8 flags of the series 5 times each (each flag was shown during the presentation phase and then 4 times during the tournament, since each team played 4 matches). All the teams won and lost twice, the position (right or left) of the winning team being random. The experimenter made sure that the children continued to pay attention during the exposure phase.

Test

After training, the prerecorded voice introduced the children to the next part of the game and a monkey appeared on the screen with a blank flag (Fig. 2d).” The following day, it is the monkeys’ turn to play “tug of war”. Oh, look! The monkey has forgotten to put colors on its flag. You can help him! You know how to make pretty flags, so help the monkey by placing the colors you want on the flag that you have in front of you. Do it now!” The same instructions were repeated for the other 5 monkeys that appeared with blank flags. In each case, the children were presented with a template consisting of either 3, 4 or 5 empty boxes, depending on the length of the flag, together with 25 squares of colors displayed in random order in front of them (see Fig. 2b). No reference to the flags seen during training was made. These implicit instructions did not require the children to make any intentional effort to retrieve information, minimizing potential age effects due to the fact that these intentional processes are known to evolve with age (e.g., Cowan, 2005). Furthermore, the behavior that the children consciously tried to produce during the test was not the same as the behavior the experimenter was trying to induce using methods of which they were unaware. Indeed, the children were asked to create "pretty flags" in the test rather than flags similar to those seen in the learning phase. They were free to select the colors of their choice, and the test instructions meant that there could be no right or wrong responses. The training items constituted primes for the subsequent generation test. These instructions made use of the criteria defined by Vinter and Perruchet (1999, 2000, 2002; see also Vinter and Detable, 2003) for preventing conscious influences from contaminating information processing during testing.[2]

When a flag was completed, the experimenter recorded the production using the numeric keypad and the colors selected by the child appeared on the monkey's flag. A pilot investigation had previously shown that it was better to ask children to build flags using color squares rather than by pressing colored keys on the keypad because, in this latter case, the children adopt strategies such as pressing adjacent keys. The colored squares selected for the flag were put back in the box. The 25 squares were thus available each time a child had to make a flag. The children were asked to produce 3 flags of 3, 4 and 5 colors respectively (random order). They then started 3 matches by pressing the "start" key (Fig. 2e).

Finally, the experiment terminated with a question phase. The experimenter asked the children some questions in the form of a debriefing: "Do you know why I asked you to take part in this game? What do you think I observed when you played in this game? Have you noticed something about the colors of the flags? About how the flags were made?" The children mainly evoked the length of the flags, the help given to the monkeys, or the colors they had seen: "You wanted to see if I can make flags." "You wanted me to make flags for the monkeys." "You wanted me to make the pandas and the monkeys play." "There were some blue, red, green, light blue and yellow colors in the flags." "There were small, medium-sized and big flags." None of the answers spontaneously evoked the presence of repetitions inside the flags or other regularities in the color sequences. None of the children explicitly linked the test episode to the training episode. These data were not analyzed further.[3]

Coding of the data

Learning of the ADJ and non-ADJ repetitions with the correct colors

We coded the production of ADJ and non-ADJ repetitions with surface-dependent features, that is the repetitions presenting the same color sequences as those seen during training. They were thus grammatical when considered in connection with the experimental groups or corresponded to the sequences seen during training in the case of the control group. These frequencies of repetitions with correct colors were calculated as a function of flag length. For example, an adjacent repetition was awarded a score of .33 in a 4-color flag (1 occurrence out of 3 possible repetitions).

Learning of the ADJ and non-ADJ repetitive structure independently of colors

We coded the production of repetitions (either adjacent or non-adjacent) with colors different from those seen in training, and computed their frequency of occurrence as explained above.

Though we used a grammar with positional constraints, the scoring procedure did not take positional constraints

---

[2] It may be worth explaining why it was not possible to include another test (such as a recognition test, for instance) following the implicit generation test. Recognition is very likely to elicit intentional information recollection processes (Gardiner & Java, 1993), with the result that such a test is a much more explicit test than an implicit generation test. Several authors have pointed out the danger of contamination between implicit and explicit tests when they are jointly administered to the same participants (Gebauer & Mackintosh, 2007; Seger et al., 2000; Shimamura, 1985). Counterbalancing the order between tests does not resolve the problem. A recognition test cannot be administered first because half of the items presented during this test are non-grammatical. However, exposing participants to errors (non-grammatical information) is known to be prejudicial to implicit learning (e.g., Baddeley & Wilson, 1994; Perruchet, et al. 2006) and this can affect performance in a subsequent test. The same reasoning holds if the generation test is presented first. Such a test inevitably leads participants to generate false (non-grammatical) items (partially or completely).

[3] This does not mean that the children could not evoke the presence of adjacent or non-adjacent repetitions. As we saw in a small number of children, if the experimenter explicitly asked "Were the flags always made of different colors, or were some colors sometimes repeated inside the flags?", they were able to recognize that some colors were sometimes repeated. The explicit knowledge that could be expressed by children in the question phase was highly dependent on the type of questions asked by the experimenter.

into account mainly because the repetitive units occupied a large part of the possible positions in the training sets (for instance, for the non-ADJ units in the non-ADJ training condition, repetitions occurred at all the positions in the 3- and 4-color flags and at 3 out of the 4 possible positions in the 5-color flags). Clearly, this material was not suited to evidence positional learning. A grammar with positional constraints was nevertheless adopted because fixed positions likely favor the detection of regularities and facilitate implicit learning processes (e.g., Mathews et al., 1989).

ANOVAS were carried out on the mean frequencies of ADJ and non-ADJ repetitions with Age (4) and Types of series (3) as between-subjects factors. The analysis concluded with comparisons between the observed and theoretical proportions based on the use of Student's *t* tests. Two methods were used to compute the theoretical proportions. The Monte-Carlo method consists in simulating the probability distribution of the tested system while taking account of various parameters (the number of colored squares, the length of the flags, and the procedure which involved selecting a color square which was not subsequently replaced), randomly sampling a large number of attempts, and then computing the frequencies of occurrence of the desired events. We also employed an analytical approach which involved computing the precise theoretical probabilities of producing a repetition (adjacent or non-adjacent) in different cases. We generated the entire set of 3-, 4- and 5-color flags that could be produced using the 25 colored squares in conditions of color-selection without replacement (e.g., for 3-color flags: 25*24*23 possibilities). On the basis of the generated set, the program then determined the number of adjacent and non-adjacent repetitions for the different flag lengths. The Monte-Carlo method and the analytical approach produced similar results. We will present the analytically computed proportions below.

Results

Learning of the ADJ and non-ADJ repetitions with the correct colors

An ANOVA was conducted with Age (4) and Type of Series (3) as between-subjects factors. The Type of Series factor had a significant effect on the production of non-ADJ repetitions, $F(2, 108) = 15.04$, $p < .01$, and ADJ repetitions, $F(2, 108) = 17.1$, $p < .01$. Planned comparisons showed that the non-ADJ group produced significantly more non-ADJ repetitions with the correct colors than the ADJ or RAND group, $p_s < .01$. The ADJ group produced significantly more ADJ repetitions with the correct colors than the other two groups, $p_s < .01$. Furthermore, Student's *t* tests showed that the production of ADJ units in the non-ADJ and control groups did not differ significantly, $t < 1$, as that of non-ADJ units in the ADJ group in comparison to the control group, $t(78) = 1.64$, $p = .10$. Figure 4 presents the production of ADJ and non-ADJ units with the correct colors as a function of the Type of Series and Age.

The observed proportions of ADJ and non-ADJ units produced in the ADJ and non-ADJ groups were compared to the theoretical distributions expected from the strong rule-based (1/1 in both cases), mild rule-based (2/1 and 3/2) and stimuli-specific (8/1 and 8/2) views (see Fig. 1). In the ADJ training condition, the participants introduced ADJ and non-ADJ units in proportions that mismatch with those drawn from the strong rule-based view, $\chi^2(1) = 12.22$, $p < .01$, and from the stimuli-specific view, $\chi^2(1) = 7.46$, $p < .01$, but did not depart from those drawn from the mild rule-based view, $\chi^2(1) = 1.28$, $p = .26$. In the non-ADJ training condition, the participants produced the two kinds of repetitive units in proportions diverging significantly from those hypothesized by the strong rule-based view $\chi^2(1) = 29.78$, $p < .01$, and mild rule-based view, $\chi^2(1) = 17.15$, $p < .01$, but not from those made within the stimuli-specific view, $\chi^2(1) = 1.28$, $p = .26$.

For the non-ADJ units, the effects of Age as well as of the Age by Type of Series interaction were not significant, $F < 1$ and $F(6, 108) = 1.02$, $p = .42$ respectively. Planned comparisons investigating the individual age groups showed that the non-ADJ group produced more non-ADJ units than the ADJ group and the RAND group at age 5 ($p < .01$ and $p < .05$ respectively), at age 6 ($p_s < .01$) and at age 7 ($p_s < .01$). At 8 years of age, children in the non-ADJ group introduced significantly more non-ADJ units with the correct colors in their flags than those in the RAND group ($p < .05$), but not more than those in the ADJ group, $F < 1$. Indeed, the ADJ group produced significantly more non-ADJ units than the RAND group ($p < .01$), as Fig. 4 illustrates. For the ADJ repetitions, the ANOVA also failed to reveal significant effects of Age, $F < 1$, or of Age by Type of Series interaction, $F(6, 108) = 1.26$, $p = .28$. Planned comparisons revealed that the ADJ group produced more ADJ units than the non-ADJ group and the RAND group at age 5 ($p_s < .01$), at age 7 ($p_s < .01$) and at age 8 ($p_s < .01$). At 6 years of age, the ADJ group significantly outperformed the RAND group ($p < .05$), but not the non-ADJ group, $F < 1$. Figure 4 shows that this failure to obtain a significant difference between the experimental groups was due to the fact that the ADJ group did not produce sufficient ADJ units, rather than to an increase of ADJ units in the non-ADJ group as compared to the RAND group ($F < 1$).

Table 2 presents the observed and theoretical proportions of ADJ and non-ADJ repetitions with the correct colors. Student's *t* tests failed to reveal significant differences between the proportions of non-ADJ units and chance in the non-ADJ and the ADJ groups, $t(39) = 1.53$, $p = .13$ and $t(39) = .19, p = .85$ respectively, whereas the RAND group produced significantly fewer non-ADJ repetitions than chance would predict, $t(39) = -2.67$, $p < .05$. Concerning the ADJ repetitions, performance in the ADJ group did not significantly differ from chance, $t(39) = 1.63$, $p = .11$, whereas the non-ADJ and RAND groups' productions were significantly lower than chance, $t(39) = -7.9$, $p < .01$ and $t(39) = -10.27$, $p < .01$ respectively.[4] Performance in the RAND group revealed a natural tendency to produce repetitions at a level significantly below chance in the absence of any external influences. This is an important finding because it means that the theoretical proportions do not provide good estimations of how children might behave in the absence of external influences such as those exerted by the grammatical series in the experimental groups.

Finally, our material allowed us to go some way toward testing whether frequency was the main determinant of the specific learning observed in the ADJ group or whether perceptual saliency also contributed. In effect, the initial bigrams of the flags exhibited the same frequency as the ADJ units in the ADJ series but did not include any salient perceptual feature such as repetition. The ADJ group introduced these frequent bigrams without repetitions in their flags in 8.4% of cases (theoretical proportion: 8.3%), with the equivalent value for ADJ units being 9.4% (theoretical proportion: 6.7%). To compare these frequencies, we computed the ratio between the observed and theoretical proportions for each variable and participant. A Student's t test failed to reveal a significant difference, $t(39) = 1.69$, $p = .11$. However, given that the theoretical proportion associated with the production of adjacent repetitions is probably overestimated (6.7%), as the performance observed in the control groups (1.5%) reveals, this result suggests that perceptual saliency may compete with statistical frequency in the learning of specific items.[5]

Production of ADJ and non-ADJ repetitive units independently of colors

An ANOVA was run on the frequencies of ADJ or non-ADJ units made of colors different from those seen during training with Age (4) and Type of Series (3) as between-subjects factors. Figure 5 illustrates the production of these repetitive units as a function of Age and Type of Series.

The Type of Series factor was not significant for these non-ADJ units, $F < 1$. By contrast, it was significant for the production of ADJ units, $F(2, 108) = 8.2$, $p < .01$, with the ADJ group performing better than the other two groups. The ANOVAs failed to reveal any Age effects, $F_s < 1$, or Age by Type of Series effects, $F_s < 1$. When the age groups were analyzed separately, the ANOVAs showed a reliable main effect of Type of Series for the ADJ units, with the ADJ group always outperforming the non-ADJ and RAND groups ($p_s < .05$), except at 7 years of age, $F_s < 1$. Figure 5 shows that the production of ADJ repetitive units seemed to be particularly high at 6 years of age. Type of Series was never significant for the non-ADJ repetitive units, whatever the age, $F_s < 1$.

The results of the Student's t tests, which compared the observed proportions of the ADJ and non-ADJ repetitive units in colors different from those shown in training with the theoretical ones, are presented in Table 3. The ADJ and non-ADJ groups produced both ADJ and non-ADJ repetitive units at chance level, $t(39) = -.74$, $p = .46$ and $t(39) = .31, p = .76$ respectively, whereas the RAND group performed at significantly below-chance level for both types of repetitive units, as did the non-ADJ group for the ADJ structure, and the ADJ group for the non-ADJ structure ($p_s < .01$).

Discussion

This study investigated whether children would develop sensitivity to the grammatical structure or to the perceptual

---

[4] Perruchet and Reber (2003, p. 129) argued that "*if trained control subjects do not perform at chance level, this means that some undetected biases are still in operation*". The below-chance production of repetitions in the random group probably expressed a natural bias against repetitions. Because the additivity assumption claims that non-specific variables have similar effects on the experimental and control groups trained in very close conditions, if repetition avoidance led to below-chance productions in the control group, it was very likely that sensitivity to repetitive structures raised performance in the experimental groups. However, this increase reached only chance level, not more. For a more direct demonstration of such learning, we computed the ratios between observed and theoretical proportions, and run an ANOVA with Type of series as a between-subjects factor. Type of series yielded significance for the non-adjacent and adjacent repetitions, $F(2, 117) = 7.05$, $p < .01$ and $F(2, 117) = 18.25$, $p < .01$ respectively. Planned comparisons showed that the non-adjacent group produced more non-adjacent repetitions than the adjacent and random groups, $F(1, 117) = 9.46$, $p < .01$ and $F(1, 117) = 11.58$, $p < .01$ respectively, as did the adjacent group for the adjacent repetitions, $F(1, 117) = 26.85$, $p < .01$ and $F(1, 117) = 27.9$, $p < .01$. Thus, there was a specific increase of adjacent or non-adjacent repetitions only in the groups where specific learning was expected.

[5] The results suggest that the frequent initial bigrams were not learned, as their mean frequency was very close to the theoretical probability. However, it is worth pointing out that the second color in these bigrams was the first color of the repeated adjacent unit (e.g., BYY). Learning the initial bigram (BY) may have been prevented, or at least slowed down, by the formation of the repeated unit (YY). It would appear to be very interesting to study this issue in greater detail because it reveals the limitations of a purely statistical view of learning.

Fig. 4 Mean frequencies of non-adjacent and adjacent repetitions produced with the correct colors in the implicit generation test as a function of age (5, 6, 7 or 8 years) and Type of series (non-ADJ, ADJ or RAND)

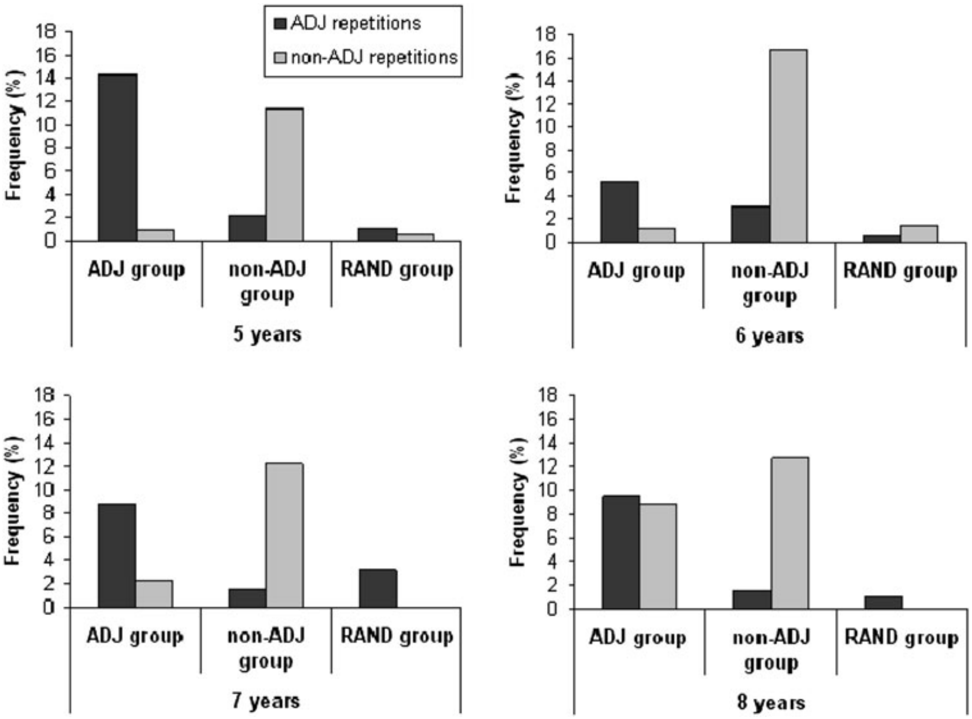


Table 2 Student's *t* test comparing observed proportions of non-ADJ and ADJ repetitions with the correct colors to theoretical ones

Units	Group	Theoretical proportions	Observed proportions	Student's *t* test
Correct non-ADJ repetitions	Non-ADJ	.093	.135	$t(39) = 1.64$, $p = .11$
	ADJ	.031	.033	$t(39) = .19$, $p = .85$
	RAND-control	.031	.015	*$t(39) = -2.67$, $p < .05$
Correct ADJ repetitions	Non-ADJ	.067	.028	*$t(39) = -4.02$, $p < .01$
	ADJ	.067	.094	$t(39) = 1.63$, $p = .11$
	RAND-control	.067	.015	*$t(39) = -10.27$, $p < .01$

* Indicates significant differences with $p < .05$ or .01

properties of the material they were exposed to in an AGL context, and whether they would be able to learn some kind of abstract relational information. Overall, the results provide support for the models of implicit learning that claim that these processes are rooted in the processing of surface features of the stimuli, and not in the abstraction of grammatical rules, and that sensitivity to relational information can develop, at least for simple relations.

The children in the experimental groups were exposed to items taken from the same artificial finite state grammar but presenting different surface features (ADJ or non-ADJ repetitions). The strong rule-based view assumes that the knowledge acquired during training refers to the grammatical structure of the material, and not to its surface features (Hypothesis 1). The ADJ and non-ADJ series differed only in terms of frequencies of salient units but not in terms of the structural rules. Consequently, this theory does not predict differences in learning between the ADJ and non-ADJ groups. Our results failed to support this view. The children preferentially learned the salient units of the series that they saw. Indeed, the ADJ and non-ADJ groups produced more ADJ and non-ADJ repetitions with the correct colors, respectively, than the other groups. This specific learning effect is quite remarkable, especially as performance in the control group revealed a clear, spontaneous tendency to introduce repetitions in the generation of sequences of colors at significantly below-chance level (for an analysis of this effect, see Witt and Vinter, in press).

A mild rule-based view claims that the knowledge acquired during an IL episode relates to the legal co-variations seen in the material, which reflect the structural rules and should be learned by frequency-independent abstraction processes (Hypothesis 2). The productions of ADJ or non-ADJ units should depend on the number of ADJ or non-ADJ rules instantiated in each series shown to the experimental groups, and should significantly differentiate

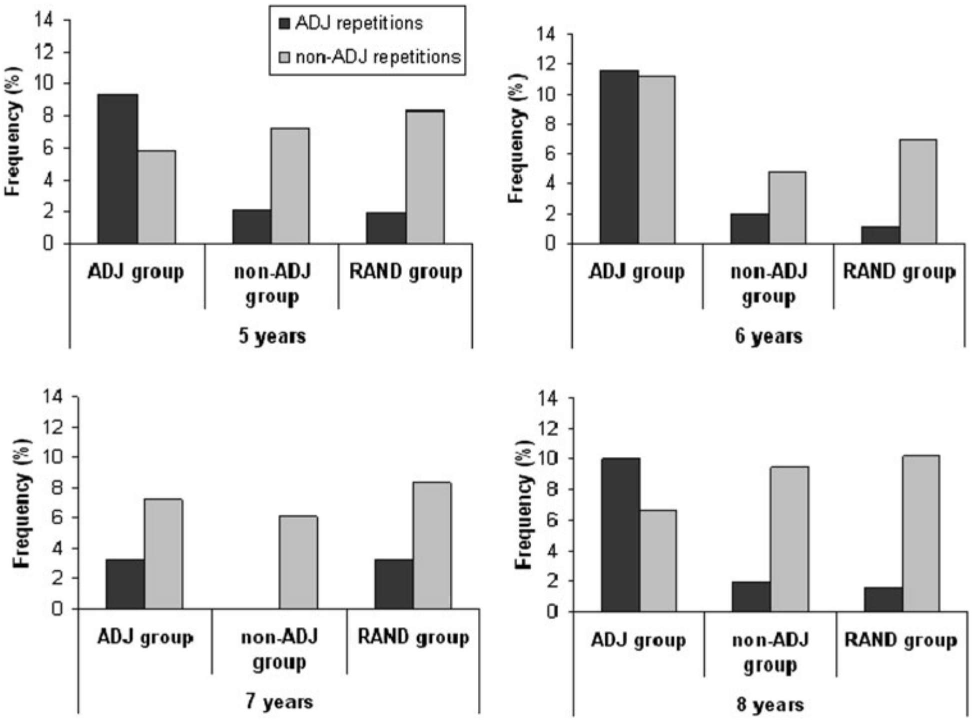


Fig. 5 Mean frequencies of non-adjacent and adjacent repetitions with incorrect colors in the implicit generation test as a function of Age (5, 6, 7 or 8 years) and Type of series (non-ADJ, ADJ or RAND)

Table 3 Student's *t* test comparing observed proportions of non-ADJ and ADJ repetitions with incorrect colors to theoretical ones

Repetitive structures	Group	Theoretical proportions	Observed proportions	Student's *t* test
Non-ADJ	Non-ADJ	.062	.066	$t(39) = .31, p = .76$
	ADJ	.124	.077	$*t(39) = -3.50, p < .01$
	RAND-control	.124	.085	$*t(39) = -3.15, p < .01$
ADJ	Non-ADJ	.100	.021	$*t(39) = -10.43, p < .01$
	ADJ	.100	.085	$t(39) = -.74, p < .46$
	RAND-control	.100	.019	$*t(39) = -12.15, p < .01$

* Indicates significant differences with $p < .05$ or .01

these groups from the control group. A larger production of ADJ and non-ADJ units was indeed observed in the ADJ and non-ADJ groups respectively, in proportions similar to those drawn from the number of rules contained in the ADJ series, but in much larger proportions than those expected from the number of rules contained in the non-ADJ series. Furthermore, the critical comparisons with the control group failed to reach significance: The ADJ group did not produce more non-ADJ units than the control group, and similarly, the production of ADJ units in the non-ADJ group did not depart from the level observed in the control group. Though this view makes predictions much closer to the observed performance than a strong rule-based view, it cannot account for the whole pattern of results, especially in the case of the non-ADJ training condition. Yet, the statistical coverage of the grammar was higher in the non-ADJ series than in the ADJ series, which should have encouraged rule abstraction in the non-ADJ training condition rather than in the ADJ one according to Poletiek and van Schijndel (2009).

The pattern of behaviors recorded in the non-ADJ training condition appears congruent with the predictions drawn from stimulus-specific processing models (Hypothesis 3). Children in the non-ADJ group produced more non-ADJ units than those in the ADJ or control group, and as many ADJ units as those in the control group. The same results appeared in the ADJ group. Whether children memorized exemplars (Brooks, 1978) or developed conscious representations of the situation as a function of the content of their attentional focus (Perruchet & Vinter, 2002), they were expected to learn the specific features of the material they were exposed to. These findings suggest that models, which support the idea of stimulus-specific processing in AGL, are as valid for children as they are for adults (Johansson, 2009). In our view, perceptual saliency and frequency were the main stimulus features that guided specific item learning in our experiment. The design of the material does not make it possible to distinguish between the roles of these two features. This issue deserves to be investigated in greater detail. However, as mentioned earlier, the mild rule-based view accounted partly for the results in the ADJ group as well. Looking at the age effects may provide more insights in this discussion.

The differences observed between age groups are indeed worth examining. The profile of results shown by the 6-year-olds in the ADJ group appears closer to the one predicted by the mild rule-based view than by the stimulus-specific view, in terms of respective proportions of ADJ and non-ADJ units produced, but not in terms of the comparisons with the control group. However, the production of adjacent repetitive units (independently of colors) recorded at this age enables to account for this profile without suspecting that the rules have been "partially" abstracted. The 6-year-olds might have been particularly sensitive to the adjacent repetition structure itself, thus causing them to generate many repetitions, regardless of color, as illustrated in Fig. 5. This high level of production of adjacent repetitive units may have prevented them to generate a higher rate of adjacent units with the correct colors. Indeed, a strong tendency to avoid the production of repetitions has been reported by a number of authors in studies of healthy children and adults asked to generate sequences of discrete events (Frith, 1972; Mittenecker, 1958; Wiegersma, 1982; Witt and Vinter, in press). This anti-repetition bias constrains the introduction of repetitions within sequences, a large production of repetitions regardless of colors preventing the production of repetitions with specific colors.

The result profile exhibited by the 8-year-olds in the ADJ group appears more controversial. It seems conform to the one predicted from a strong version of the rule-based view, suggesting that these children have abstracted the rules in this training condition, while they would have developed a sensitivity to the specific surface features of the material in the other condition. These findings could indicate that these older children had recourse to both rule-based and stimulus-specific processing during the IL episode, as some studies conducted with adults have concluded (Knowlton & Squire, 1996; Meulemans & Van der Linden, 1997; Tunney & Altmann, 2001). Rule-based processing would be facilitated in response to exposure to the ADJ series, possibly because they implemented a low-complexity, first-order rule. This asymmetrical result profile in the 8-year-olds would thus seem to be due both to the simplicity of the material to be processed (Tunney & Altmann, 2001) and to the fact that older children are better able to make use of abstraction processes than their younger counterparts (e.g., Case, 1985; Fischer, 1980). The greater procedural flexibility shown by children around this age (Karmiloff-Smith, 1992; Picard & Vinter, 1999; Vinter & Marot, 2007) may also explain why rule-based and stimulus-specific processing can conjointly operate during the learning episode. Further developmental studies are needed to confirm this suggestion that abstraction of rules in older children can be observed in the case of simple material.

As was expected from the models discussed here, the children demonstrated that they were able to learn relational information, which constitutes some form of abstract information. This shows that abstraction is not necessarily indicative of rule formation and rule use, as cogently argued by Redington and Chater (2002, p. 4): "… surface-independence and rule-based knowledge are orthogonal concepts." Our results indicate that children generalized the adjacent repetitive configuration in the ADJ group, as shown by the significant production of adjacent repetitive units with ungrammatical colors, whereas they were unable to generalize the non-ADJ repetitive relation in the non-ADJ group. This asymmetrical profile of results can be accounted by the different models. Abstracting a first-order dependency rule is surely less complex than abstracting a second-order dependency rule. The similarity between XX and YY is greater than between XaX and YbY. The direct perception of the overall structure of repetition is probably facilitated when two elements need to be associated rather than three. An alternative explanation relates to the much higher variance of the non-ADJ patterns (B_B, R_R, J_J) in the non-ADJ series than of the ADJ patterns (YY, GG) in the ADJ series. On the one hand, a high variance could make non-ADJ repetitions less salient, preventing their generalization. On the other hand, a high variance reduces the frequency of specific patterns and following Smith et al. (1992), this should be beneficial to rule use. Children failed to generalize the relational information while they processed items with high variance. This indicates that rule abstraction did not operate despite favorable conditions.

To conclude, our overall findings provide more support for models that consider the IL processes to be anchored in the specific properties of the material shown to children during training. Of course, further studies will be required in order to check whether factors such as the duration of learning, the complexity of the material to be learned, the number of items presented at test, or the degree of statistical coverage of the grammar contained in the training items modify the results in children in the same way as they seem to do in adults. Research into IL in children is still at an early stage. Our results also testify to the importance of including trained control groups in an AGL paradigm, especially when generation tasks are used at test. The control groups that took part in our experiment revealed that chance response levels do not constitute an acceptable estimation of how children might behave in a generation task such as ours in the absence of any prior influence or sensitization such as that resulting from exposure to structured material. Finally, the use of generation tasks in AGL may appear problematic since participants can build sequences regardless of what they learned, or at least, under the influences of spontaneous biases of production that can interact with the expression of behavioral adaptations to the training material. We acknowledge that these spontaneous biases elicited by the generation task conduced to a "not straightforward" learning demonstration, calling for further experiments to confirm or refine the present results. However, the novel methodology used in our experiment appears most promising for opening new interesting issues.

Acknowledgments The authors are very grateful to Stéphane Argon and Laurent Bergerot, who designed and programmed the video game. We thank Pierre Perruchet and Paul Molin very much for their precious help in the programming of the algorithms for computing these theoretical proportions, and Pierre Perruchet for his very constructive comments on earlier versions of the manuscript. We are also grateful to the Inspection Académique de Dijon that gave us the opportunity to run our experiments inside some kindergarten and primary grade schools. This research was supported by a grant from the Conseil Régional de Bourgogne.

References

Baddeley, A., & Wilson, B. A. (1994). When implicit learning fails: Amnesia and the problem of error elimination. *Neuropsychologia, 32*, 53–68.

Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 169–221). Hillsdale, NJ: Erlbaum.

Brooks, L. R., & Vokey, J. R. (1991). Abstract analogies and abstracted grammars: Comments on Reber (1989) and Mathews et al. (1989). *Journal of Experimental Psychology: General, 120*, 316–323.

Carr, T. H., & Curran, T. (1994). Cognitive factors in learning about structured sequences: Applications to syntax. *Studies in Second Language Acquisition, 16*, 205–230.

Case, R. (1985). *Intellectual development: Birth to adulthood*. New York: Academic Press.

Chandler, S. (1993). Are rules and modules really necessary for explaining language? *Journal of Psycholinguistic Research, 22*, 593–606.

Cowan, N. (2005). *Working memory capacity*. New York: Psychology Press.

Endress, A. D., Dehaene-Lambertz, G., & Mehler, J. (2007). Perceptual constraints and the learnability of simple grammars. *Cognition, 105*, 577–614.

Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review, 87*, 477–531.

Frith, U. (1972). Cognitive mechanisms in autism: Experiments with color and tone sequence production. *Journal of Autism and Childhood Schizophrenia, 2*, 160–173.

Gallistel, C. (1990). *The organization of learning*. Cambridge, MA: MIT Press.

Gallistel, C. (2000). The replacement of general-purpose learning models with adaptively specialized learning modules. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (2nd ed., pp. 1179–1191). Cambridge, MA: MIT Press.

Gardiner, J. M., & Java, R. I. (1993). Recognising and remembering. In A. F. Collins, S. E. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 163–188). Hove, UK: Lawrence Erlbaum Associates.

Gebauer, G. F., & Mackintosh, N. J. (2007). Psychometric intelligence dissociates implicit and explicit learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 33*(1), 34–54.

Gomez, R. L., & Gerken, L. A. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition, 70*(2), 109–135.

Jiménez, L. (Ed). (2003). *Attention and implicit learning*. Amsterdam: John Benjamins Publishing Company.

Johansson, T. (2009). Strengthening the case for stimulus-specificity in artificial grammar learning. *Experimental Psychology, 56*, 188–197.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Knowlton, B. J., & Squire, L. R. (1996). Artificial grammar learning depends on implicit acquisition of both abstract and exemplar-specific information. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 22*, 169–181.

Krist, H., Fieberg, E. L., & Wilkening, F. (1993). Intuitive physics in action and judgment: The development of knowledge about projectile motion. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 19*, 952–966.

Ling, J., & Blades, M. (1996). Incidental recall of color information by children and adults. *Applied Cognitive Psychology, 10*, 141–150.

Ling, J., & Blades, M. (2002). Further evidence for incidental encoding of color by children and adults. *British Journal of Developmental Psychology, 20*, 537–544.

Manza, L., & Reber, A. S. (1997). Representing artificial grammar: Transfer across stimulus forms and modalities. In D. Berry (Ed.), *How implicit is implicit learning* (pp. 73–106). Oxford: Oxford University Press.

Mathews, R. C., Buss, R. R., Stanley, W. B., Blanchards-Fields, F., Cho, J. R., & Druhan, B. (1989). Role of implicit and explicit process in learning from examples: A synergistic effect. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 15*(6), 1083–1100.

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.

Mittenecker, E. (1958). Analysis of sequences of "chance" action. *Zeitschrift fur Experimentelle und Angewandte Psychologie, 5*, 45–60.

Monaghan, P., & Rowson, C. (2008). The effect of repetition and similarity on sequence learning. *Memory and Cognition, 36*, 1509–1514.

Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning in real world context: The case of orthographic regularities. *Journal of Experimental Psychology: General, 130*, 401–426.

Patel, H., Blades, M., & Andrade, J. (1999). Children's incidental recall of color information. *British Journal of Developmental Psychology, 41A*, 797–828.

Perruchet, P. (1994). Defining the knowledge units of a synthetic language: Comment on Vokey and Brooks. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 20*, 223–228.

Perruchet, P., & Reber, R. (2003). Why untrained control groups provide invalid baselines: A reply to Dienes & Altmann. *The Quarterly Journal of Experimental Psychology, 56A*, 125–130.

Perruchet, P., Rey, A., Hivert, E., & Pacton, S. (2006). Do distractors interfere with memory for study pairs in associative recognition? *Memory & Cognition, 34*, 1046–1054.

Perruchet, P., & Vinter, A. (1998). Learning and development. The implicit knowledge assumption reconsidered. In M. Stadler & P. Frensch (Eds.), *Handbook of implicit learning* (pp. 495–531). Thousands Oaks: Sage publications.

Perruchet, P., & Vinter, A. (2002). The Self-Organizing Consciousness (Target paper). *Behavioral and Brain Sciences, 25*, 297–330.

Picard, D., & Vinter, A. (1999). Representational flexibility in children's drawings: Effects of age and verbal instructions. *British Journal of Developmental Psychology, 17*, 605–622.

Poletiek, F. H., & van Schijndel, T. J. P. (2009). Stimulus set size and statistical coverage of the grammar in artificial grammar learning. *Psychonomic Bulletin & Review, 16*, 1058–1064.

Pothos, E. M. (2007). Theories of artificial grammar learning. *Psychological Bulletin, 133*, 227–244.

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal learning and Verbal Behavior, 6*, 855–863.

Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*, 219–235.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. New York: Oxford University Press.

Reber, A. S., & Lewis, S. (1977). Implicit learning: An analysis of the form and structure of a body of tacit knowledge. *Cognition, 5*, 333–361.

Redington, M., & Chater, N. (2002). Knowledge representation and transfer in artificial grammar learning. In R. A. French & A. Cleeremans (Eds.), *Implicit learning*. Hove, United Kingdom: Psychology Press.

Seger, C. A., Prabhakaran, V., Poldrack, R. A., & Gabrieli, J. D. E. (2000). Neural activity differs between explicit and implicit learning of artificial grammar strings: An fMRI study. *Psychology, 28*, 283–292.

Shimamura, A. P. (1985). Problems with the finding of stochastic independence as evidence for multiple memory systems. *Journal of Experimental Psychology: Human Perception and Performance, 11*(1), 1–13.

Smith, E. E., Langston, C., & Nisbett, R. E. (1992). The case of rules in reasoning. *Cognitive science, 16*(1), 1–40.

Stadler, M., & Frensch, P. (Eds.) (1998). *Handbook of implicit learning*. Thousands Oaks: Sage Publications.

Tunney, R. J., & Altmann, G. T. M. (2001). Two modes of transfer in artificial grammar learning. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 27*(3), 614–639.

Vinter, A., & Detable, C. (2003). Implicit learning in children and adolescents with mental retardation. *American Journal on Mental Retardation, 108*, 94–107.

Vinter, A., & Marot, V. (2007). The development of context sensitivity in children's graphic copying strategies. *Developmental Psychology, 43*, 94–110.

Vinter, A., & Perruchet, P. (1999). Isolating unconscious influences: The neutral parameter procedure. *Quarterly Journal of Experimental Psychology, 52A*, 857–875.

Vinter, A., & Perruchet, P. (2000). Implicit learning in children is not related to age: Evidence from drawing behavior. *Child Development, 71*, 1223–1240.

Vinter, A., & Perruchet, P. (2002). Implicit motor learning through observational training in adults and children. *Memory & Cognition, 30*, 256–261.

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.

Wiegersma, S. (1982). A control theory of sequential response production. *Psychological Research, 44*, 175–188.

Witt, A., & Vinter, A. (in press). Learning implicitly to produce avoided behaviours. *The Quarterly Journal of Experimental Psychology*. doi:10.1080/1747.0218.2010.543