CrossMark

Discussion

What baboons can (not) tell us about natural language grammars

Fenna H. Poletiek [a,b,*], Hartmut Fitz [b], Bruno R. Bocanegra [a]

[a] Leiden University, Leiden Institute for Brain and Cognition, Leiden, The Netherlands
[b] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

A B S T R A C T

Rey et al. (2012) present data from a study with baboons that they interpret in support of the idea that center-embedded structures in human language have their origin in low level memory mechanisms and associative learning. Critically, the authors claim that the baboons showed a behavioral preference that is consistent with center-embedded sequences over other types of sequences. We argue that the baboons' response patterns suggest that two mechanisms are involved: first, they can be trained to associate a particular response with a particular stimulus, and, second, when faced with two conditioned stimuli in a row, they respond to the most recent one first, copying behavior they had been rewarded for during training. Although Rey et al. (2012) 'experiment shows that the baboons' behavior is driven by low level mechanisms, it is not clear how the animal behavior reported, bears on the phenomenon of Center Embedded structures in human syntax. Hence, (1) natural language syntax may indeed have been shaped by low level mechanisms, and (2) the baboons' behavior is driven by low level stimulus response learning, as Rey et al. propose. But is the second evidence for the first?

We will discuss in what ways this study can and cannot give evidential value for explaining the origin of Center Embedded recursion in human grammar. More generally, their study provokes an interesting reflection on the use of animal studies in order to understand features of the human linguistic system.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The title of Rey, Perruchet, and Fagot (2012) (hereafter RPF)'s paper summarizes their hypothesis: "Centre embedded structures are a by-product of associative learning and working memory constraints: Evidence from baboons". The authors took on the important challenge to investigate the foundations of complex syntax in low level cognitive mechanisms, by looking at non human primates. The present discussion is not about the first part of the proposal,

that "CE structures in human syntax are a by-product of associative learning and working memory constraints". It addresses the second part: "Evidence from baboons". Are RPFs' animal data evidence for this claim? We analyze the methodology and the logic of the RPF study, and discuss the implications about the origins of natural language constructions, accordingly. More generally, we explore the possible contribution of interpreting animal behavior in terms of human linguistic cognition, for understanding human language form.

2. The study

RPF report on data from non human primates (baboons) that, according to the authors, support the view that

hierarchical Center Embedded structures in natural language have emerged as a result of low level mechanisms. The critical evidence supporting this possibility is the baboons' "preference" for a "Center Embedded consistent (CE) pattern" of a pair of subsequent responses over a "non-embedded" (NE) sequential pattern that is not consistent with such pattern.

The animals were trained intensively on sequential associations between pairs of stimuli a and b, belonging to category A and category B, respectively, in multiple training blocks. The stimuli were shapes, and the associations between A- and B-shapes was arbitrary. In the last training block, the animals were first presented a stimulus "a" on one screen. Next, they were presented a screen with two b-stimuli, from which they select one. The animal is rewarded if it selects the b associated with the a displayed in the previous screen. In a subsequent test phase (Test 2, Fig. 1, p. 182), the last training block procedure is repeated, but now two "A"-screens, each with an a-stimulus (first a_1, then a_2), are presented, before the animal can respond to b-stimuli. After the second A-screen (displaying a_2), one B-screen with three b's is shown (b_1, b_2, b_3): b_2 being the associate of a_2, presented on the last A-screen, b_1 being the associate of a_1 presented on the first A-screen, and b_3 being a distractor, i.e. non-associated with any a presented in the two preceding screens. The baboon responds to the b's on the B-screen as follows: After having 'tapped' one of the b's, it disappears from the screen. Then nothing happens until the animal taps a second b. The baboons are rewarded after having tapped any two of the three b's on the display, in any order. In summary, the animals see one screen with an A stimulus (a_1), followed by a screen with another A stimulus (a_2), followed by one screen with three b-stimuli on random locations: b_1 (associated with a_1), b_2 (associated with a_2), and b_3 not associated to either of the a's.

The results show that, overall, the *first* b-stimulus tapped by the baboons is most often b_2 (110 times per baboon on average), which is associated to the a_2 on the screen presented just before the B-screen shows up. The next most frequent *first* response is b_1 (73 times per baboon). b_3 (unassociated to any of the a's presented just before) is chosen as a first response 53 times, on average. The *second* b selected was an associate (either b_1 or b_2) in a slight majority of cases: 55%. 45% of *second* b-responses were the non-associated b_3. When the initial response was b_3, the second response (being necessarily one of the two associated b's left over on the screen) was *equally* often b_1 as b_2.

Importantly, RPF report the frequencies of baboons' choices in terms of units of *two* consecutive responses selections (Fig. 1). For example, a b_2 followed by b_1 "b_2b_1" is a unit, b_2b_3 is one et cetera; in total six units (i.e. ordered pairs) of two consecutive responses are distinguished and counted. The b_2b_1 unit is labeled "Center Embedded structure (CE)", or "Center Embedded consistent structure"; likewise, the b_1b_2 unit is labeled the "Non Embedded structure (NE)" or "Non Embedded consistent structure". The other paired b-responses have no special label. Statistical tests show that baboons (1) "produce significantly more CE-structures than

NE-structures", that (2) after selecting a b_2, baboons had "a preference for b_1, the last element of a CE structure" over a non-associated b_3. RPF do not report explicitly two other contrasts in their data, that might yet help to understand the baboons behavior in the task: First, given an initial selection of either b_1 or b_2, the baboons were equally likely to respond to the other associated element (b_2 or b_1) as they were to the non-associated one b_3. Second, units containing a non-associated response b_3 (i.e. b_3b_1, b_3b_2, b_1b_3 or b_2b_3), were chosen more often than any of the CE- or NE-units.

The baboons' more frequent b_2b_1 response pairs as compared to b_1b_2 responses, are interpreted to "reflect their preference for center-embedded structure", and to be "inconsistent with an NE structure". The other four types of response pairs containing a b_3 (distractor), and therefore *not* looking like any equivalent syntactic rule in language, are not involved in the main comparative analysis. These type of response pairs were chosen most often.

What do these results tell us about hierarchical syntax in human languages? Many languages, including English, allow for the formation of CE structures (e.g., $a_1a_2a_3\ldots$ $b_3b_2b_1$) whereas other types of sequential ordering of dependencies, like non-embedded (NE) structures, corresponding to crossed dependencies in natural grammars (e.g., $a_1a_2a_3\ldots b_1b_2b_3\ldots$), are typologically rare. RPF propose that "the production of CE structures in baboons and humans could be the by-product of associative mechanisms and working memory constraints". CE structures may not reflect a uniquely human capacity to process the long distance dependencies characteristic of recursive CE structures (however see Hauser, Chomsky & Fitch (2002) and Fitch & Hauser's (2004) study that failed to find learning of CE structure in cotton top tamarins), but might have originated in low level cognitive mechanisms, present in baboons.

We acknowledge the importance of RPF's attempts to specify low level memory and associative learning processes that may have favored CE structures in language. Without disputing or advocating their proposal, the aim of the present discussion is to analyze the value and limitations of animal data like the ones presented by RPF, for explaining characteristics of language form.

3. Description of what the baboons do

Crucially, to appreciate RPF's claim, we need to know what is meant by interpretative labels "preference", "production", and "CE-structures" and "the baboons produce/prefer CE-structures/CE consistent structures". RPF's reasoning from the baboons responses to the conclusion is made in three steps. As a first reasoning step, the analyses of the paired responses showing $b_2b_1 > b_1b_2$, the baboons are said to "display a *preference* for *CE (consistent) structures*". Secondly, this preference is argued to be clearly *not* caused by an *actual* preference for a grammatical structure ("our findings do not imply that the baboons possess the innate computational device that has been postulated for humans"). So, the frequent b_2b_1b choices are claimed to be a *by-product* of elementary associative mechanisms and working memory processes, rather than a preference

for a sequence. According to the third reasoning step, CE structures in human language production could *have their origins* in the same (associative learning and working memory) mechanisms that are also resourced by the baboons to perform the experimental task.

Below we explore the conditions that studies on animal behavior should satisfy, to be informative about CE-syntax in human language. In particular, we will discuss what mechanisms best explain the baboons' responses in RPF's task. Second, we compare the baboons' task, and the task in which humans produce sentences with CE dependencies. Third, we investigate what *mechanisms* could explain *both* the data in RPF's study and in humans producing CE sentences. In this way, we reflect on the possibilities and limits of animal studies for understanding features of human grammars.

4. What the baboons do

What are the baboons doing? Clearly, we agree with RPF that "the baboons do not possess any innate computational device" that makes them prefer CE structures rather than NE structures. The question remains: *do* the baboons *prefer* CE like structures over NE like structures, in any similar way human languages 'prefer' CE over NE sequencing? Did the baboons preferably order a series of associated responses in a 'mirror' fashion (b_n, b_{n-1}, b_{n-2}, b_{n-3}...) rather than in a 'copy' fashion (b_1, b_2, b_3,...,b_n) in the task they performed? The observations RPF propose to support that possibility are that b_2b_1 response pairs outnumbered b_1b_2. At face value, the latter result may speak for CE consistency. However, the baboons more often choosing a pair containing a non-associate, than associates only, speaks *against* both CE and NE consistency. Crucially, thus, what does result the high occurrence of b_2b_1 imply? In order to sort out the implications of this observation for the occurrence of CE rather than NE structures in human languages, we need to know what mechanism is underlying the higher number of b_2b_1 response pairs over b_1b_2 pairs in this study. Was it a '*preference*' for sequencing two associated responses in a CE fashion rather than in a NE fashion? Or was it any other mechanisms that can produce sequences of responses in the specific task carried out by the baboons?

Logically, if the high number of b_2b_1 responses are indicative of a preference for a sequential structure ("b_2-and-then-b_1"), then the two B-choices should be interdependent. Hence, b_2 should be chosen first *with the objective to select it first, ahead of b_1*. Alternatively, b_2 could also be responded to first by the animals for other reasons than its relative position to b_1. In that case, choosing b_2 first should not be predictive of the response that comes next. If separate factors explain independently the first and second tapping, a description of the results in terms of the two independent factors would be a more parsimonious account of the results. The direct inference route from the "CE-consistency" of the baboons response pairs, to the CE structure of language, would be blocked, and another reasoning would be needed to link the two domains.

The dependence hypothesis of the b_2b_1 response pair, can be tested in RPF's data. If RPF's interpretation is correct, an initial b_2 choice should be followed more often by a b_1 choice, than an initial b_1 choice is followed by a b_2 choice. If both transitions occur equally often, however, the relative positions of the b_1 to b_2 has played no role in the baboons' preference for a response pair. In Table 1 a Chi-square test for independence of the first and second choice, is performed.

A Chi-square test with Yates' continuity correction shows that the proportion of type of second responses was independent of whether the first response had been consistent with a CE (b_2 first) or NE structure (b_1 first), $X^2(1, N = 2005) = 0.55$, $p = 0.458$. After choosing either an initial b_2 or b_1, the baboons show a slight bias in favor of a second associate choice, rather than the distractor. If the baboons actually prefer the CE sequence over NE, this bias should be larger for completing b_2-first choices than for completing b_1-first choices. This bias is exactly the same for both CE and NE consistent first choices, though.

If no preference for CE or NE is reflected in the responses, what could explain the large number of b_2b_1 observed? Critically, the data reported by RPF can be explained by two tendencies which affect the base-rate of responding without affecting the contingencies between responses: first, a tendency to respond to recent information before responding to more distant information (recency bias), and secondly, a tendency to choose an associated item rather than a non-associated item (association bias). Adding up these two independent tendencies makes the $b_2b_1 > b_1b_2$ contrast emerge in an analysis of the response frequencies as ordered pairs.

Summarizing, for the baboons, tapping the b that is associated with the a displayed on the last screen encountered, is a continuation of the behavior they had been rewarded for during the last part of the training-phase. This shows that the animals were able to associate two stimuli in a rewarded training procedure. Hence, although RPF clearly show associative learning of individual A and B stimuli in the baboons, the data do not show a CE orderings of multiple associated stimuli. The data show no "preference for any of two *sequence structures*" of responses. Rather, the baboons responded to recent information before responding to more distant information. Secondly, and independently, the animals were more sensitive to associated items than to non-associated items. Notice that despite this latter capacity, however, response *pairs* containing a *non-associated* item were more frequently chosen than pairs without any such distractor, overall, suggesting that the task quite often elicited response patterns

Table 1
Frequencies of second choice after the first CE- and the first NE-consistent response (Rey et al., 2012).

Second response	First response		
	CE structure consistent	NE structure consistent	
Structure consistent completion	658	453	1111
Unrelated distractor b_3	545	349	894
	1203	802	2005

unrelated to any associations learned, and therefore obviously also unrelated to preferences regarding the sequencing of associations.

5. Is a CE-structure-consistent response the production of a CE-structure?

Even if the CE-consistent pattern in RPF's data does not emerge from a preferred sequencing tendency, the high number of associated b_2b_1 units from among the units with associates, might be related to language structures via mechanisms that may sub serve both tasks. To explore this, we need to specify the common features of (1) RPF's experimental task, and (2) the task of a language user producing a sequence of verbs after a sequence of nouns, in a multiple clauses sentence, as in *the dog(a_1) the man(a_2) walks(b_2) barks(b_1)*.

Both tasks are similar in that they require the pairwise association of elements, each from one of two categories, and lining them up in a sequence. However, the task goals largely differ. The language producer's goal is to line up words he or she intends to associate with each other, in order to convey meaning. Since meaning varies with word ordering – a crucial feature of syntax-successful word ordering validly represents the intended meaning (choosing *the girls the boy <u>kisses eats</u>*; versus *the girl the boy <u>eats kisses</u>*). In the animal study, different sequences of tappings did not differentially affect the reward in RPF's design. Simulating this crucial feature of human language in the type of experiment RPF carried out, does not seem unfeasible, however. For example, a different type of reward might be given for different response orderings; and one might observe whether the baboons select specific orderings to selectively control the reward.

A second difference between the baboons' and the language user's task, is that constructing a CE sentence crucially assumes an arbitrary but equal number of A's and B's in a sentence. The baboons got a reward after a fixed number of two responses. Possibly, choosing freely the number of responses would have affected their response ordering. Due to this fixed number of possible responses, the experiment leaves untested the multiple possible ordering preferences that are at odds with CE structures (e.g. $b_2b_1b_3$). Experiments that aim to determine whether an animal's sequential behavior is characterized by CE structure should therefore test the animal's ability to tally and check off the number of A-responses against the number of B-responses (Fitch & Hauser, 2004; Perruchet & Rey, 2005; De Vries, Monaghan, Knecht, & Zwitserlood, 2008), and their selective preference for a mirror ordering of any number of responses.

Taken together, both human language producers and RPF's baboons are faced with a task requiring them to line up responses associated with a series of stimuli presented just before. This basic similarity sets the minimal condition for transferring what the baboons do in their task, to what humans do in language production. But is it enough? Of course, experimentation, inherently, is a controlled simulation of natural conditions that trades off validity against controllability. The validity of experimental results is therefore a matter of degree. The constraints of RPF's experiment, including the reward structure, the transition from the training to the test phase, the options offered, the degrees of freedom of choice, and the lack of effects following from different response orderings, distinguish it in essential respects from the conditions under which humans produce sentences. Considering, in addition, that the CE-consistency of the data emerged partly because of the format used analyze the data (counting pairs), and that no *preference* for CE-sequences over any other ordering was demonstrated, the question whether a mechanism at work in the baboons' performance may have put pressure on the occurrence of CE grammar in human syntax, remains an interesting, yet open question.

6. Discussion

RPF's study shows that baboons can learn associations between two categories (A and B) of abstract visual stimuli, and tended to respond first to the B-item that was associated with the most recently presented A-stimulus. However, from all response pairs, most contained one non-associated B (a distractor). Importantly, the large number of responses to the distractor, questions the relevance of these responses for assessing sequential structures of associated elements in language. Moreover, RPF's data do not show that the first choice statistically affected the choice for the second one, suggesting that the frequent CE-consistent response pair (b_2b_1) were not due to a preference to order responses in that particular way. Taken together, the baboons' behavior may not warrant claims concerning the occurrence of CE sequential structure in human language.

An important question RPF's research raises is: if such a preference were demonstrated in animals, what would that imply about mirror patterns in human syntax? Such data would suggest that these mechanisms are really low-level. An interesting continuation to RPF's research project would then be to establish whether, in humans, the same general mechanisms lead to preferences for mirror ordering behavior, independently of language.

The effect of low-level cognitive constraints on language is now a main focus of exploration in studies with human participants. Christiansen and Chater (2008) argue in a review that general information processing characteristics of the human brain have shaped the form of language. MacDonald (2013) discusses studies on language production, showing that language form is shaped by general motor-, and memory-related constraint. Lai and Poletiek (2011) and Poletiek and Lai (2012) show that processing CE structure, is largely facilitated by prior non-linguistic knowledge about associations. Interestingly, RPF report a similarity between response orderings produced by baboons and ordering structure in language. However, it remains to be shown whether these results are relevant for explaining the production of CE orderings in human languages. One of the problems we pointed out, that illustrate the challenge of relating animal data to human cognition, is that producing a response sequence, is not the same as sequencing your responses.

Acknowledgements

References

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences, 31*, 489–558.

De Vries, M. H., Monaghan, P., Knecht, S., & Zwitserlood, P. (2008). Syntactic structure and artificial grammar learning: The learnability of embedded hierarchical structures. *Cognition, 107*, 763–774.

Fitch, W. T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science, 303*, 377–380.

Lai, J., & Poletiek, F. H. (2011). The impact of adjacent-dependencies and staged-input on the learnability of center embedded hierarchical structures. *Cognition, 118*, 265–273.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology* (4). http://dx.doi.org/10.3389/fpsyg.2013.00226.

Perruchet, P., & Rey, A. (2005). Does the mastery of center-embedded structures distinguish humans from non humans? *Psychonomic Bulletin & Review, 12*, 307–313.

Poletiek, F. H., & Lai, J. (2012). How semantic biases in simple adjacencies affect learning a complex structure with non-adjacencies in AGL: A statistical account. *Philosophical Transactions of the Royal Society B, 367*, 2046–2054. http://dx.doi.org/10.1098/rstb.2012.0100.

Rey, A., Perruchet, P., & Fagot, J. (2012). Centre-embedded structures are a by-product of associative learning and working memory constraints: Evidence from baboons. *Cognition, 123*, 180–184.