Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates?

Pierre Perruchet *, Arnaud Rey

University of Bourgogne, Dijon, France

Running Head: Learning hierarchical structures

Word count: 4267 (main text + references)

* Corresponding author:
Pierre Perruchet
Université de Bourgogne
LEAD/CNRS
Pole AAFE, Esplanade Erasme
21000 Dijon, France

email: pierre.perruchet@u-bourgogne.fr

Abstract

In a recent *Science* paper, Fitch and Hauser (2004; hereafter, F&H) claimed to have demonstrated that Cotton-top Tamarins fail to learn an artificial language produced by a Phrase Structure Grammar (PSG, Chomsky, 1957) generating center-embedded sentences, while adult humans easily learn such a language. We report an experiment replicating the results of F&H in humans, but also showing that participants learned the language without exploiting in any way the center-embedded structure. When the procedure was modified to make the processing of this structure mandatory, participants no longer showed evidence of learning. We propose a simple interpretation for the difference in performance observed in F&H's task between humans and Tamarins, and argue that, beyond the specific drawbacks inherent to F&H's study, researching the source of the inability of nonhuman primates to master language within a framework built around the Chomsky's hierarchy of grammars is a conceptual dead-end.

Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates?

Identifying the specific language abilities that separate humans from nonhuman primates has been the topic of innumerable speculations. In a recent *Science* paper, Fitch and Hauser (2004; hereafter F&H) argued that the hierarchy of grammars of increasing generative power described by Chomsky (e.g., 1957) provides the key for a response. At the lowest level of complexity are the finite state grammars (FSG), which generate sequences by concatenating a set of elements (states) while following pre-specified transitional probabilities. The extant literature, as F&H point out, indicates that nonhuman primates can master such grammars. However, according to Chomsky, human language use requires the mastery of the next level in the complexity hierarchy, termed the "phrase structure grammar" (PSG). "In addition to concatenating items like a FSG, a PSG can embed strings within other strings, thus creating complex hierarchical structures and long-distance dependencies" (F&H, p.378). An instance of embedding in English is "the rat the cat ate stole the cheese", in which one relative clause ("the cat ate") is nested within the sentence ("the rat stole the cheese"). The aim of the F&H paper was to show that although the abilities to master a PSG are available to all normal humans, they are not available to monkeys.

In their experimental demonstration, F&H used a particular PSG, termed A^nB^n, where n={2,3}. This grammar generates center-embedded constructions, such as that represented in Figure 1 for n=3. The A and B elements were drawn within separate sets of eight CV syllables, and were further distinguished by their acoustic characteristics. The A syllables were spoken by a female and the B syllables by a male, so that the two classes of syllables differed by voice pitch, quality, and other particularities of the voice sources.

Participants (undergraduate students on the one hand, and cotton-top Tamarins on the other) first heard a set of sentences following the patterns AABB or AAABBB. In the subsequent test phase, they heard novel sentences, half following the same grammar (A^nB^n) and half following a finite-state grammar (AB^n), which generated either ABAB or ABABAB sentences. Students were asked to state whether the pattern of each novel sound was the same as or different from the pattern heard during the familiarization phase. They scored 85% correct on this discrimination task. The performance of the Tamarins was assessed through their visual orientation towards the loudspeaker, an increase in looking rate being taken as indicative that the sounds were perceived as different. Interestingly, Tamarins displayed an equivalent rate of looking to strings that violated the rules of the grammar and to strings that were consistent with those rules. Obviously, their failure to selectively look at the non-consistent strings could be due to multiple causes, such as a perceptual inability to discriminate the acoustic properties of A and B syllables. In order to eliminate alternative interpretations, F&H inverted the two grammars for other groups of subjects. It turned out that Tamarins trained with ABAB or ABABAB sequences showed a significant increase in looking to the AABB or AAABBB strings when they were displayed during the test phase. Because this inverted task involved the same perceptual abilities as the first one, F&H inferred that the failure of Tamarins trained with the PSG provided a demonstration of their inability to master this class of grammar.

Are F&H's conclusions compelling? Regarding their demonstration that students are able to master a PSG, F&H remained cautious, noting that "limited output from a PSG can always be approximated by a more complicated FSG" (p. 378). But they pursued: "However, failure to master a grammar (as demonstrated by a failure to distinguish grammatical from ungrammatical strings) can be empirically confirmed" (p.378). It is notable that there is a logical contradiction between these two statements. If it is not possible to demonstrate that the achievement in a specific task *t* implies the mastery of a grammar *g*, then how could it be possible to conclude that the failure in *t* attests that *g* cannot be mastered? The starting point of the present reappraisal is that *the failure of Tamarins in the F&H task is relevant with regard to their alleged inability to master a PSG only if it can be asserted with a reasonable confidence that the achievement of humans in the same task attests for their mastery of the PSG*. Our main objective is to show that this condition was not fulfilled, because the performance of humans in the task used by F&H was unrelated to their discovery of the center-embedded structure of the language.

F&H repeatedly claimed that the crucial aspect at hand in their experiment concerns the discovery of the hierarchical structure inherent in the center-embedded pattern of the auditory strings. This sustained emphasis may lead the reader to believe that the pairing between A1 and B1, A2 and B2, and A3 and B3 (referring to the labels used in Figure 1) is mandatory, as is the pairing between "the rat" and "stole the cheese" and "the cat" and "ate" in the sentence above. Only these pairwise relationships ensure us that AAABBB sequences are center-embedded (although the point is not explicit in the A^nB^n notation). Now, these relationships, if present in the experimental materials (F&H mentioned nowhere that syllables were paired), were, in any case, unnecessary for grammatical discrimination. F&H acknowledged this point when they claimed that the achievement in their task "requires some way to recognize a correspondence between either the groups formed by the As and Bs (e.g., counting) or between specific As and corresponding Bs (e.g., long-distance dependencies)." (p.378) Whether a sentence with the same number of A and B syllables, irrespective of syllable-to-syllable matching, instantiates a genuine PSG appears to be questionable.[1] But even if one takes the point as given, it is worth noting that counting was not even necessary to achieve perfect discrimination. Indeed, there was no test sequence following the patterns AABBB or AAABB (i.e. in which the number of A differs from the number of B) in the F&H final test. As a consequence, there is no way to claim that participants had detected that there were the same number of A and B syllables in the strings displayed during the familiarization phase. A parsimonious interpretation may be that human participants simply discriminated the cases where there was one female-to-male voice transition (AABB or AAABBB) from the cases in which there were two or three consecutive alternations (ABAB or ABABAB). In keeping with our argument above, since there is no evidence that the achievement of humans attests for their mastery of a PSG, the assertion that the failure of monkeys in the same task is proof for their inability to master a PSG is clearly unwarranted.

Even if respecting Occam's razor principle favors our alternative interpretation, it may be argued however that the possibility that participants used a simple strategy is not proof that they did not use a more complex one. It remains possible that participants have learned the language as a center-embedded structure, thus making long-distance associations between, for example, the first and the last syllables of each sequence. In order to test this possibility, we conducted a new experiment borrowing many of the materials and procedures from F&H, with the following modification. The sentences implemented a genuine center-embedded grammar, in addition to the structure used in F&H. The A syllables and the B syllables were paired in a consistent way for each subject during the training phase. For instance, assuming that *ba* was paired with *gu* for a participant, if *ba* occurred as A1, then *gu* occurred as B1, if *ba* occurred as A2, then *gu* occurred as B2, and if *ba* occurred as A3, then *gu* occurred as B3. During the test phases, four categories of items were presented for rating, resulting from the crossing of two binary dimensions. Test sentences were consistent or inconsistent with the acoustic pattern heard during the familiarisation phase (i.e., two or three high pitched syllables, followed by two or three low pitched syllables - a test close to the one implemented in F&H[2]), and consistent or inconsistent with the center-embedded structure. The final design is shown in Table 1. If participants are sensitive to the center-embedded structure, performances should exhibit a main effect of, or an interaction with this factor.

A second prediction involved a comparison of the effects of the acoustic pattern observed when n=2 and n=3. If participants processed the material as a center-embedded structure, the performance would be lower when n is larger, on the grounds that the complexity of center-embedded sentences increases as a function of the depth of embeddings (e.g., Blaubergs & Braine, 1974; see also Ellefson, 2002, for a study on artificial grammar learning with visual letters). By contrast, the simple strategy we have proposed above leads to inverted predictions. Indeed, if participants simply discriminated between one female/male transition on the one hand, and several successive alternations on the other hand, the discrimination would be as easier as the number of successive alternations became larger.

Method

Participants. A total of 32 undergraduate students from the University of Bourgogne in Dijon, France, participated in the experiment in partial fulfillment of a course requirement. All subjects were native French speakers.

Materials. The syllables composing the sentences were those used in F&H, with only minor changes (*yo* was replaced by *ro*; *wu* by *vu*; and *pa* by *sa*) intended to increase syllable discrimination in French. Possible A syllables were {ba di ro tu la mi no vu}, and possible B syllables were {sa li mo nu ka bi do gu}. The speech was synthesized using the MBROLA speech synthesizer (http://tcts.fpms.ac.be/synthesis/, Dutoit, Pagel, Pierret, Bataille, & Van Der Vrecken, 1996) with the FR4 diphone database. The resulting WAV file was played through headphones connected to a PC computer using Windows Media Player.

The duration of each syllable was 450 ms (with 150 ms for consonants and 300 ms for vowels), a value that approximated that used by F&H (because this information was not reported, we estimated syllable duration from the sample files provided on the Science web site). The A syllables were set at a fundamental frequency of 240 Hz, and the B syllables at 80 Hz. Although the A and B syllables differed only along the pitch dimension (whereas F&H

used samples of female and male voices, which differed along various dimensions such as phonetic identity), the pitch variation was very salient. The sentences were obtained by combining the syllables, without any silent pauses between syllables.

Procedure. Participants were tested individually in a sound-attenuated room. They were told that they would listen to sequences of sounds for three minutes and that they would be asked to answer questions about the sounds at the end of the presentation. They were asked to avoid engaging in analytic, problem solving processes. Thirty-two sentences of the language were then presented, half of them comprising two pairs of syllables, and half three pairs (F&H presented 30 sentences; We selected the nearest multiple of eight, in order to display equally often the eight A and B syllables). The two or three A syllables composing the first half of each sentence were drawn pseudo-randomly, with the following constraints: (1) a given syllable never occurred twice in the same sentence and (2) a given pair of adjacent syllables did not occur more than three times across the whole list, hence keeping at a low level the transitional probabilities between adjacent syllables. Finally (3) over all the sentences, each syllable was presented four times in first location and four times in second location. Once the A syllables were drawn, the B syllables were generated according to the center-embedded grammar shown in Figure 1. The pairing between the A and B syllables was arbitrary, and differed for each participant. As a consequence, over the entire sample of participants, the grammatical status of any sentence was independent from its surface form, hence preventing the possibility that judgments were affected by a priori perceptual biases. The sentences were pseudo-randomly ordered for each participant, and separated by a silent period of 3400 ms.

After familiarization, participants were told that they would be presented with a set of novel auditory strings, and that they would have to judge, for each one, if the pattern was the same as or different from the pattern of the strings heard previously. The Experimenter noted participants' verbal responses. There were 16 test sentences, which were designed as shown in Table 1. Half of them were consistent with the acoustic pattern and half were not; likewise, half of them were consistent with the center-embedded structure and half were not; and finally, half of them comprised two pairs of syllables, and half comprised three pairs. The sentences inconsistent with the acoustic pattern were generated as in F&H, and the sentences inconsistent with the center-embedded structure were obtained by scrambling the B syllables composing the legal second half of each sentence. The test sentences were given in random order, with the order differing for each participant. For the sake of illustration, Appendix A provides the sentences to which Participant #1 was exposed during the study phase and the test phase.

Finally, participants heard a list of 16 trisyllabic strings, in the same conditions as those used during the study and the test phases. Eight strings were AAA and eight strings were BBB. Each syllable occurred once in each of the three possible locations within the strings. Participants were asked to write down on a sheet of paper the sounds that they perceived. This task was devised to check whether participants perceived the sounds correctly, in order to rule out the possibility that a failure to exploit the pairwise associations between syllables can be accounted for by a poor perceptual identification.

<u>Results</u>

<u>Spelling test</u>. Most (93.4 %) of the syllables were spelled correctly in the final dictation (i.e., were consistent with French grapho-phonological transcription rules). Among the errors, 37.2% consisted in the transcription of "ka" as "ga", and 8.8% in the transcription of "vu" as "bu". These misspellings were generally consistent within participants. Most of the other misspellings occurred once or twice, and many of them consisted of minor confusions of vowels (e.g. transcription of /o/ as "ou" (/u/) or "on" (/õ/)). None of these errors could be

detrimental for syllable discrimination within the set of syllables used in the experiment (e.g. the syllables "ga" and "bu", and the vowels "ou" and "on", were not part of the materials).

Pattern recognition test. Figure 2 shows the percentage of "Different" responses collected during the test phase, according to whether the sounds were a violation of, or were consistent with the sound heard during the study phase. There was a striking effect of the acoustic pattern, which is similar to the effect reported by F&H (expressed as a percentage of correct responses, participants obtained a score of 82.62%, while F&H reported 85%), whereas there appears to be no effect of the embedding. An ANOVA was run with the variations in Acoustic pattern, the variations in Grammatical status, and the Number of syllables composing test items as within-subject factors. There was a main effect of Acoustic pattern ($F(1,31) = 118.77$, $p <.0001$), while the effect of Grammaticality was not significant ($F(1,31) = 0.97$, $p = .33$). The Acoustic pattern by Grammaticality interaction was not significant either ($F(1,31) = 0.42$, $p = .52$) , indicating that the detection of violations in the acoustic pattern was independent of the consistency of the strings with the center-embedded structure displayed during the familiarisation phase.

There was no main effect for the number of syllables composing the strings ($F(1,31) = 0.93$, $p = .34$), but there was a reliable Acoustic pattern by Number of syllables interaction ($F(1,31) = 10.36$, $p = .003$). The pattern of results is displayed in Figure 3. It appears that the difference between the number of "Different" responses given to inconsistent and consistent strings was larger when n=3 than when n=2, indicating that performance was better when the sequences were longer. As explained above, this result provides complementary evidence that participants did not process the material as a center-embedded structure. The pattern of results is consistent with the idea that participants simply discriminated between one high-pitch/low-pitch transition on the one hand, and several successive alternations on the other hand.

No other interaction was significant.

Discussion

The present results provide compelling evidence against the idea that human participants processed the acoustic variations in the F&H language as a center-embedded structure. First, we failed to reveal any sensitivity to the center-embedded pattern provided by the pairing of A and B syllables, despite our relatively large sample of subjects (*N*=32). Second, participants' sensitivity to changes in the acoustic pattern was better when the strings were longer, a pattern that should have been the reverse had the material been processed as a center-embedded structure. These data are consistent with the hypothesis that human participants performed the test as a simple perceptual discrimination task. As a consequence, the contrasted results of humans and Tamarins reported by F&H on a similar task cannot be attributed to the absence in the latter of a specific ability to process center-embeddings.

A question immediately arises: If the discrimination was trivial, why should Tamarins have failed in this task? Addressing this question is important, because the failure of Tamarins could be construed as indirect evidence that humans engaged high-level, sophisticated problem solving strategies lying beyond monkey abilities. This conclusion is not straightforward however. Accordingly, F&H took care to eliminate some alternative interpretations, such as differences in perceptual abilities. However, it remains that humans and monkeys were submitted to quite different tests. Students were asked to discriminate the strings consistent and inconsistent with regard to the sound pattern heard previously, and they presumably tuned their response criterion in order to share their responses roughly equally among "same" and "different". By contrast, Tamarins presumably turned towards the loudspeaker only if the sounds emitted by the loudspeaker were biologically significant. This difference deeply undermines a direct comparison between the performances of humans and Tamarins. But why did Tamarins turn towards the loudspeaker when they heard AAABBB after being familiarized with ABABAB, and not the reverse? Although we are limited to speculations, one hypothesis is the following.[3] As any reader can check from listening to the sounds available on the Science web site, the AAABBB strings sound much more like natural human language than the succession of syllables alternately spelled out by the female and male voices that composed the ABABAB strings. This may explain why Tamarins selectively oriented towards the loudspeakers when they heard AAABBB after having been familiarized with the other structure. The reverse did not occur, possibly because the "novelty" introduced by ABABAB presented no potential interest (e.g., the new sounds could not cue the possible presence of humans).

Our paper's title asks the question: Does the mastery of center-embedded structures distinguish humans from nonhuman primates? Our conclusion at this point is that the *Science* paper by Fitch and Hauser (2004) does not provide a response, and hence that any speculation based on this paper (e.g., Friederici, 2004) may be premature. To go a step further, let us now address a more general issue: beyond the methodological drawbacks that call into question the soundness of the F&H study, is it still worthwhile to pursue experimental investigations about the Tamarins' abilities to master PSG grammars? A positive response would entail that such grammars are (1) commonly mastered by humans and (2) actually crucial for describing human language structures. We believe that neither of these conditions is met.

Regarding the first condition, F&H seemingly took for granted that the abilities needed to learn a hierarchical structure "are available to all normal humans" (p.378). A PSG, they claim, "is trivially easy for humans to learn" (p.378). These claims are misleading.[4] Let us add simply one more embedding to our initial sentence, and we get the oft-cited "the rat the cat the dog chased ate stole the cheese", which is unintelligible to most English speakers, as Miller & Chomsky (1963) noted themselves about a very similar example (see also Blaubergs & Braine, 1974). The literature on self-embedding is entirely devoted to accounting for why self-embedded structures are not manageable whenever the depth of embeddings exceeds one or two, even when semantic biases are available (e.g., Gibson & Thomas, 1999). Experimental studies using artificial languages do not authorize greater optimism. The experiment reported above did not reveal any learning, even though relationships were between specific tokens instead of exemplars of syntactic classes. Conway, Ellefson, and Christiansen (2003) also failed to obtain above chance performance in a task devised to reveal learning of a center-embedded structure with auditory materials (although under certain conditions, they provided positive results with visually displayed strings; see also Ellefson, 2002). Earlier experimental studies (e.g. Cleeremans, 1993; Santelmann & Jusczyk, 1998) showed that learning nonadjacent dependencies, which is a prerequisite for detecting a center-embedded structure, is only possible when the distance between the to-be-associated elements does not exceed a very small number of elements. Other studies (Gomez, 2002; Newport & Aslin, 2004; Onnis, Monaghan, Chater & Richmond, under revision; Perruchet, Tyler, Galland, & Peereman, under revision) suggests that learning nonadjacent dependencies occurs only in conditions where there is some extraneous reason to associate the relevant elements (e.g. when the prosodic pattern makes the relations salient).

A second condition for giving sense to the exploration of the monkeys' ability to process hierarchical structures is that such structures are essential for human language. This indeed has been largely accepted after Chomsky’s (1957) initial statement that human languages can not be described by finite-state grammars (FSG). Recent works in linguistic however tend to reverse this view and to consider that finite state machines can process much of language (e.g., Klavans & Resnik, 1996). Let us assume, for the sake of argument, that F&H were right in claiming that humans easily master a PSG, while Tamarins can only master a FSG. This would explain at best why Tamarins fail to process a sentence such as "the rat the cat ate stole the cheese",[5] but this would not account for their failure to process "the cat ate the rat that stole the cheese" or "the dog chased the cat that ate the rat that stole the cheese". Indeed, these right-branching structures are normally generated by a FSG, which is easily mastered by the Tamarins, according to F&H. Clearly, Tamarins do not fail only to master center-embedded sentences. They fail to master even the most elementary linguistic utterance.

In summary, we have provided evidence that the conclusions reached by F&H about the inability of nonhuman primates to master PSG grammars are unwarranted. But beyond the limits of this specific study, we doubt that the hierarchy of grammars described by Chomsky can be of much help to understanding why animals do not communicate as humans do.

References

Blaubergs, M.S., & Braine, M.D.S. (1974). Short-term memory limitations on decoding self-embedded sentences. Journal of Experimental Psychology, 102, 745-748.

Chomsky, N. (1957). Syntactic structures. Mouton, The Hague.

Cleeremans, A. (1993). Mechanisms of implicit learning : A connectionist model of sequence processing. MIT press: Bradford Books.

Christiansen, M.H., & Chater, N. (1999). Toward a connectionist model of recursion in human linguistic performance. Cognitive Science, 23, 157-205.

Conway, C.M., Ellefson, M.R., & Christiansen, M.H. (2003). When less is less and when less is more: Starting small with staged input. In Proceedings of the 25th Annual Conference of the Cognitive Science Society. (pp.270-275). Mahwah, NJ: Lawrence Erlbaum.

Dutoit, T., Pagel, N., Pierret, F., Bataille, O., & Van Der Vrecken, O. (1996). The MBROLA Project : Towards a Set of High-Quality Speech Synthesizers Free of Use for Non-Commercial Purposes. Proceedings of the ICSLP, Philadelphia, vol. 3, 1393-1396.

Ellefson, M.R. (2002). The difficulty of learning complex structure: A Comparative study of knowledge acquisition. Unpublished doctoral dissertation, Southern Illinois University, Carbondale.

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

Friederici, A.D. (2004). Processing local transitions versus long-distance syntactic hierarchies. Trends in Cognitive Sciences, 8, 245-247.

Gibson, E., & Thomas (1999). Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. Language and Cognitive Processes, 14, 225-248.

Gomez, R. (2002). Variability and detection of invariant structure. Psychological Science, 13, 431-436.

Klavans, J.L., & Resnik, P. (Eds.) (1996). The balancing act: Combining symbolic and statistical approaches to language. Cambridge, MA: MIT Press.

Miller, G., & Chomsky, N. (1963). Finitary models of language users. In R. Luce, R. Bush, & E. Galenter (Eds.). Handbook of mathematical psychology, (vol. 2, pp. 419-493), New-York, Wiley.

Morgan, J.L, & Newport, E.L. (1981). The role of constituent structure in the induction of an artificial language. Journal of Verbal Learning and Verbal Behavior, 20, 67-85.

Morgan, J.L., Meier, R.P., & Newport, E. (1989). Structural packaging in the input to language learning: Contributions of prosodic and morphological marking of phrases to the acquisition of language. Cognitive Psychology, 19, 498-550.

Newport, E.L., & Aslin, R.N. (2004). Learning at a distance: I. Statistical learning of non-adjacent dependencies. Cognitive Psychology, 48, 127-162.

Onnis, L., Monaghan, P., Chater, N., and Richmond (in revision). Phonology impacts segmentation and generalisation in speech processing.

Pereira, F.C.N., & Wright, R.N. (1997). Finite-state approximation of phrase-structure grammars. In E. Roche & Y. Schabes (Eds.). Finite state language processing (pp. 149-173). MIT Press, Cambridge.

Perruchet, P., Tyler, M.D., Galland, N., & Peereman, R. (under revision). Learning non-adjacent dependencies: No need for algebraic-like computations.

Santelmann, L.M., & Jusczyk, P.W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. Cognition, 69, 105-134.

APPENDIX A

Study sentences		Test sentences	
1 balanosadogu	17 vutubagulinu	1 mirodikabimo	(C)
2 nodivunukasa	18 nomidikamosa	2 badirokagubi	(I)
3 larobido	19 romimobi	3 labagudo	(C)
4 lanosado	20 tudikali	4 tumimoli	(C)
5 minotulisamo	21 rotulibi	5 tunolisa	(I)
6 dimimoka	22 noromimobisa	6 vudiladokanu	(C)
7 roladobi	23 vumibagumonu	7 miladidomoka	(I)
8 novumimonusa	24 bavuladonugu	8 robabigu	(I)
9 badivunukagu	25 mibanosagumo	9 baturobiligu	(C)
10 mitubagulimo	26 milanosadomo	10 norotubisali	(I)
11 vuromimobinu	27 tubaguli	11 dinosaka	(C)
12 divunuka	28 tunosali	12 rovunubi	(C)
13 tuladoli	29 ditulika	13 dimikamo	(I)
14 lavunudo	30 vubarobigunu	14 nolatulidosa	(C)
15 banovunusagu	31 robagubi	15 lavudonu	(I)
16 ladikado	32 dirobika	16 vutulalinudo	(I)

The table displays the study and the test sentences heard by Participant #1. For this participant, "ba" was consistently paired with "gu", in such a way that if "ba" occurred as A3, then "gu" occurred as B3 (e.g., Study sentence #1), if "ba" occurred as A2, then "gu" occurred as B2 (e.g., Study sentence #27), and if "ba" occurred as A1, then "gu" occurred as B1 (e.g., Study sentence #10). Likewise, "la" was paired with "do", and so on. The test sentences were either consistent (C) or inconsistent (I) with the grammar. For instance, Test sentence #2 was inconsistent, because "ba" was paired with "bi" instead of being paired with "gu" (more generally, the second half of this sentence should have been "bikagu" instead of "kagubi").

Authors Note

This work was supported by grants from the Centre National de la Recherche Scientifique (CNRS, UMR 5022), and the University of Bourgogne. The authors thank Nadege Doignon for her assistance in running the experiment. Correspondence concerning this article should be addressed to Pierre Perruchet (pierre.perruchet@u-bourgogne.fr), Université de Bourgogne, LEAD/CNRS, Pole AAFE, Esplanade Erasme, 21000 Dijon, France.

Footnotes

1- In fact, Chomsky (1957) included among the productions that could not be described by a finite state grammar the AAABBB strings in which there is no term-to-term correspondence between A and B elements and hence, that can be processed by simple counting. However (1) he distinguished these latter from the center-embedded structures, which F&H designated as their main target throughout their paper and (2) it is doubtful that such structures exist in natural language, and in fact the instance provided by Chomsky (IF..THEN) is better described as a genuine center-embedded construction (Christiansen & Chater, 1999).

2- Our test differed slightly, because the specific syllables instantiating A and B classes were left unchanged in the test strings that were inconsistent with the acoustic pattern. Thus inconsistent test strings were, for n=3, **A**A**A**B**B**B instead of **A**B**A**B**A**B (with bold and underlined characters indicating high pitched syllables). Had the syllables been changed, the factorial comparison of the two structures (pitch and center-embedding) would have not been possible. This departure could have made the discrimination of the acoustic pattern more difficult in our situation than in F&H's one. However, this possibility was not detrimental for our objective, and in fact, to anticipate, there was no sizeable difference between the performance of our and F&H's participants on this aspect.

3- Mark Liberman proposes an alternative hypothesis, see http://itre.cis.upenn.edu/~myl/languagelog/archives/000355.html

4- Commenting on their own results in humans, F&H noted: "These data are consistent with other experimental findings that humans can learn a PSG and appear to prefer phrase-structured input". Now, the empirical studies to which they refer to support their claim (Morgan & Newport, 1981; Morgan, Meier & Newport, 1989) used an artificial language generated by a Finite-State Grammar (FSG) and not a PSG (Morgan et al. indeed used the expression "Phrase-structured input", but in a sense unconnected to Chomsky (1957)'s "Phrase

Structure Grammar"). Moreover, these papers were aimed at showing that even for these simpler grammars, extraneous cues such as prosodic markers were necessary to acquire the syntax. They concluded that their result "suggests that human language learning capacities, in the adult and in the child, are limited in the amount and complexity of data that can be handled" (Morgan et al., 1989, p.546).

5- Even this rather trivial consequence is not ascertained. Indeed, it rests again on the postulate that humans use a PSG to process this kind of sentences. This postulate has no firm ground. Indeed, given the low performance on center-embedded structures and the limited importance of those structures in extant languages, much more parsimonious accounts of the processing of such sentences become possible, for instance those relying on finite-state grammar approximation (e.g. Pereira & Wright, 1997) or simple recurrent networks (Christiansen & Chater, 1999).

Table 1

Structure of the Strings Used During the Test Phase (Bold and Underlined Characters = High-Pitch; Normal Characters = Low Pitch)

			Acoustic Pattern (Pitch variation)	
			Violation	Consistent
Grammatical Structure (Center-Embedding)	Violation	*n*=2	**A1** A2 **B1** B2	**A1** **A2** B1 B2
		n=3	**A1** A2 **A3** B2 **B1** B3	**A1** **A2** **A3** B2 B1 B3
	Consistent	*n*=2	**A1** A2 **B2** B1	**A1** **A2** B2 B1
		n=3	**A1** A2 **A3** B3 **B2** B1	**A1** **A2** **A3** B3 B2 B1

Figure Captions

Figure 1. Schematic diagram of a PSG generating center-embedded sentences

Figure 2. Percentage of "Different" responses given by participants, as a function of the well-formedness of the test strings with regard to (1) their acoustic pattern and (2) their grammatical structure (i.e., center-embedded).

Figure 3. Percentage of "Different" responses given by participants as a function of the well-formedness of the test strings with regard to their acoustical pattern, according to the number of pairs of syllables composing the strings.

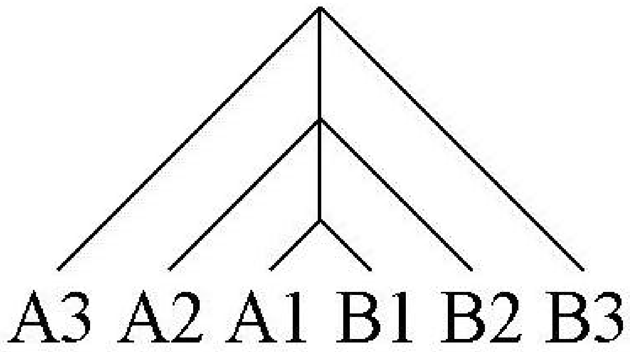

FIGURE 1

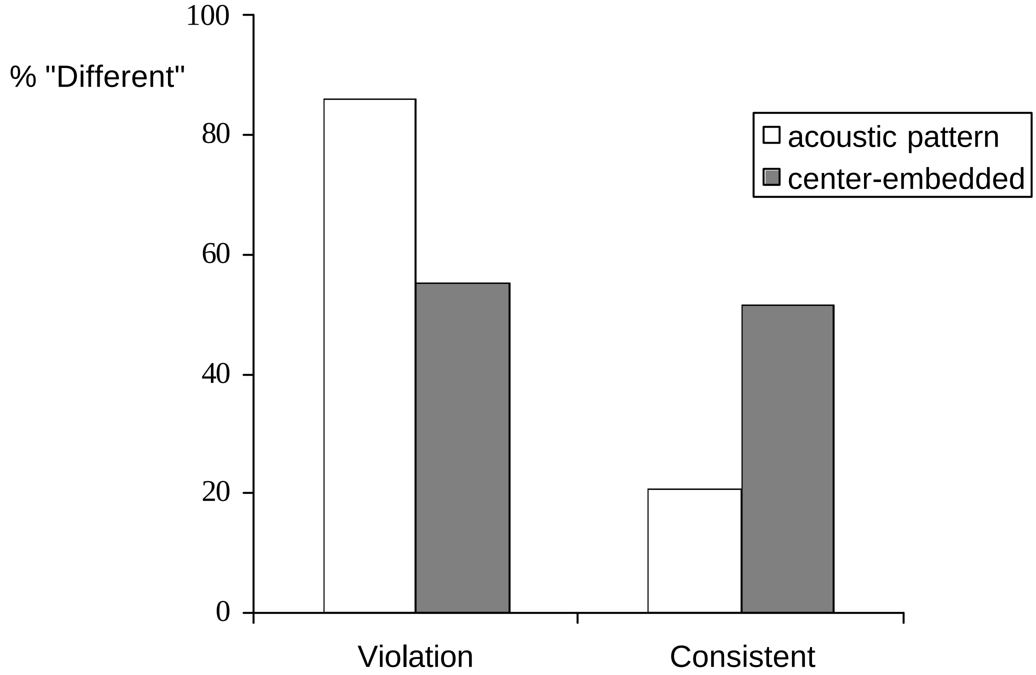
100
% "Different"
80
60
40
20
0
acoustic pattern
center-embedded
Violation
Consistent

FIGURE 2

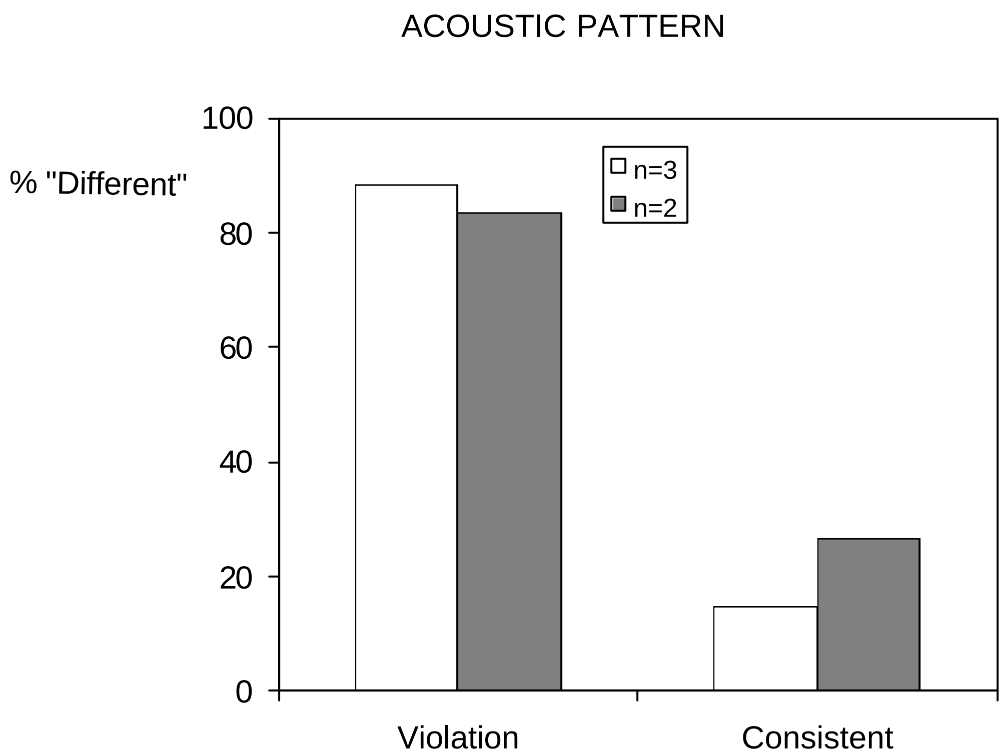
ACOUSTIC PATTERN
% "Different"
100
80
60
40
20
0
n=3
n=2
Violation
Consistent

FIGURE 3