

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

PIERRE PERRUCHET and ARNAUD REY
University of Bourgogne, Dijon, France

In a recent *Science* article, 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 (Chomsky, 1957) generating center-embedded sentences, whereas adult humans easily learn such a language. We report an experiment replicating the results of F&H in humans but also showing that subjects 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, the subjects 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 in F&H's study, researching the source of the inability of nonhuman primates to master language within a framework built around Chomsky's hierarchy of grammars is a conceptual dead end.

Identifying the specific language abilities that separate humans from nonhuman primates has been the topic of innumerable speculations. In a recent *Science* article, 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 (FSGs), which generate sequences by concatenating a set of elements (states) while following prespecified transitional probabilities. The extant literature, as F&H pointed 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 an 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 F&H's article 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 (2004) 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 pattern AABB or AAABBB. In the subsequent test phase, they heard novel sentences, half following the same grammar (A^nB^n) and half following an FSG (AB^n), which generated either ABAB or ABABAB sentences. The 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 toward the loudspeaker, an increase in looking rate being taken as indicative that the sounds were perceived as different. Interestingly, the tamarins displayed an equivalent rate of looking to strings that violated the rules of the grammar and looking to strings that were consistent with those rules. Obviously, their failure to selectively look at the nonconsistent 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 (2004) 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 the AAABBB strings when they were displayed during the test phase. Because

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 P. Perruchet, Université de Bourgogne, LEAD/CNRS, Pole AAFE, Esplanade Erasme, 21000 Dijon, France (e-mail: pierre.perruchet@u-bourgogne.fr).

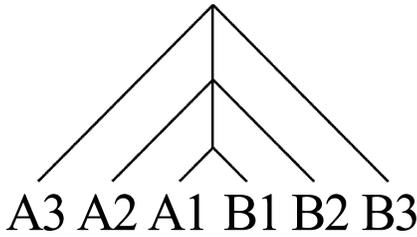


Figure 1. Schematic diagram of a phase structure grammar generating center-embedded sentences.

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 (2004) 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 continued, "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 achievement in a specific task t implies the mastery of a grammar g , how could it be possible to conclude that 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 reasonable confidence that the achievement of humans in the same task attests to 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 (2004) repeatedly claimed that the crucial aspect at hand in their experiment concerned 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 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 pattern AABBB or AAABB (i.e., in which the number of As differed from the number of Bs) in F&H's final test. As a consequence, there is no way to claim that the subjects 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 subjects simply discriminated the cases in which 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 to their mastery of a PSG, the assertion that the failure of monkeys in the same task is proof of their inability to master a PSG is clearly unwarranted.

Even if respecting Occam's razor principle favors our alternative interpretation, it may be argued that the possibility that subjects used a simple strategy is not proof that they did not use a more complex one. It remains possible that subjects 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 (2004), 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 subject, if *ba* occurred as A1, *gu* occurred as B1, if *ba* occurred as A2, *gu* occurred as B2, and if *ba* occurred as A3, *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 familiarization 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&H²) and consistent or inconsistent with the center-embedded structure. The final design is shown in Table 1. If subjects 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 subjects processed the material as a center-embedded structure, the performance would be lower

Table 1
Structure of the Strings Used During the Test Phase

| Grammatical Structure (Center-Embedding) | <i>n</i> | Acoustic Pattern (Pitch Variation) | |
|---|----------|--|--|
| | | Violation | Consistent |
| Violation | 2 | <u>A1</u> A2 B1 B2 | <u>A1</u> <u>A2</u> B1 B2 |
| | 3 | <u>A1</u> A2 <u>A3</u> B2 B1 B3 | <u>A1</u> <u>A2</u> <u>A3</u> B2 B1 B3 |
| Consistent | 2 | <u>A1</u> A2 B2 B1 | <u>A1</u> <u>A2</u> B2 B1 |
| | 3 | <u>A1</u> A2 <u>A3</u> B3 B2 B1 | <u>A1</u> <u>A2</u> <u>A3</u> B3 B2 B1 |

Note—Bold and underlined characters = high pitch; normal characters = low pitch.

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., Blaugers & 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 subjects simply discriminated between one female/male transition, on the one hand, and several successive alternations, on the other hand, the discrimination would be easier as the number of successive alternations became larger.

METHOD

Subjects

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 the subjects were native French speakers.

Materials

The syllables composing the sentences were those used in F&H (2004), 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*, and *vu*, and possible B syllables were *sa*, *li*, *mo*, *nu*, *ka*, *bi*, *do*, and *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 msec (with 150 msec for consonants and 300 msec for vowels), a value that approximated that used by F&H (2004; 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 the 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

The subjects were tested individually in a sound-attenuated room. They were told that they would listen to sequences of sounds for 3 min 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, 2004, presented 30 sentences; we selected the nearest multiple of eight, in order to dis-

play equally often the eight A and B syllables). The two or three A syllables composing the first half of each sentence were drawn pseudorandomly, 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, thus keeping at a low level the transitional probabilities between adjacent syllables. Finally, (3) over all the sentences, each syllable was presented four times in the first location and four times in the 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 the B syllables was arbitrary and differed for each subject. As a consequence, over the entire sample of subjects, the grammatical status of any sentence was independent from its surface form, thus preventing the possibility that judgments were affected by a priori perceptual biases. The sentences were pseudorandomly ordered for each subject and were separated by a silent period of 3,400 msec.

After familiarization, the subjects were told that they would be presented with a set of novel auditory strings and that they would have to judge, for each one, whether the pattern was the same as or different from the pattern of the strings heard previously. The experimenter noted the subjects' 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 (2004), 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 subject. For the sake of illustration, Appendix A provides the sentences to which Subject 1 was exposed during the study phase and the test phase.

Finally, the subjects 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 8 strings were BBB. Each syllable occurred once in each of the three possible locations within the strings. The subjects were asked to write down on a sheet of paper the sounds that they perceived. This task was devised to check whether the subjects perceived the sounds correctly, in order to rule out the possibility that a failure to exploit the pairwise associations between syllables could be accounted for by poor perceptual identification.

RESULTS

Spelling Test

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 subjects. 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 acoustic pattern, which is similar to the effect reported by F&H (2004; expressed as a percentage of correct responses, the subjects obtained a score of 82.62%, whereas F&H reported 85%), whereas there appears to have been no effect of embedding. An analysis of variance was run, with variations in acoustic pattern, variations in grammatical status, and number of syllables composing the test items as within-subjects factors. There was a main effect of acoustic pattern [$F(1,31) = 118.77$, $p < .0001$], whereas the effect of grammaticality was not significant [$F(1,31) = 0.97$, $p = .33$]. The acoustic pattern \times 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 familiarization 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 \times 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 was explained above, this result provides complementary evidence that the subjects did not process the material as a center-embedded structure. The pattern of results is consistent with the idea that the subjects 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 subjects processed the acoustic variations in the F&H (2004) 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, the subjects' sensitivity to changes in the acoustic pattern was better when the strings

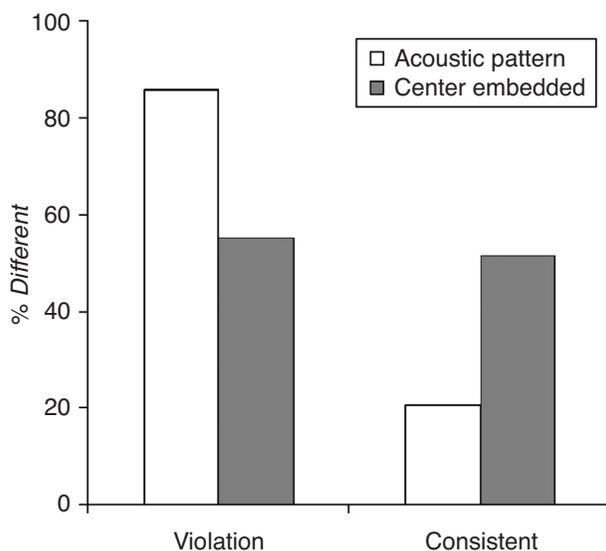


Figure 2. Percentage of *different* responses given by subjects, as a function of the well-formedness of the test strings with regard to (1) their acoustic pattern and (2) their grammatical structure (i.e., center embedded).

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 subjects 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 the tamarins have failed in this task? Addressing this question is important, because the failure of the 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 (2004) took care to eliminate some alternative interpretations, such as differences in perceptual abilities. However, it remains that the humans and the monkeys were submitted to quite different tests. The 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, the tamarins presumably turned toward 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 the tamarins turn toward 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 Sci-

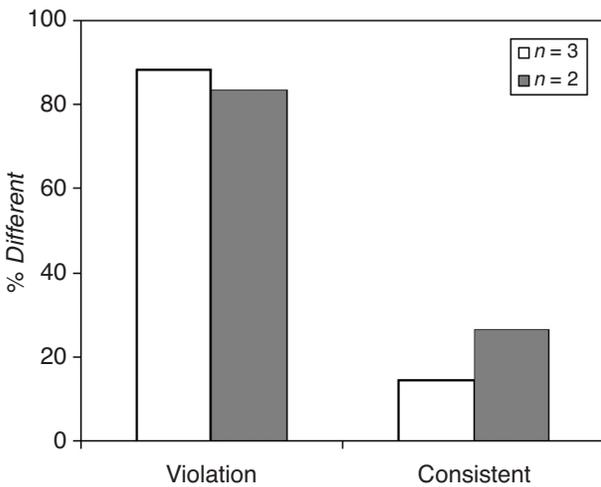


Figure 3. Percentage of *different* responses given by subjects, as a function of the well-formedness of the test strings with regard to their acoustic pattern, according to the number of pairs of syllables composing the strings.

ence Web site, the AAABBB strings sound much more like natural human language than does the succession of syllables spelled out alternately by the female and male voices that composed the ABABAB strings. This may explain why the tamarins selectively oriented toward the loudspeaker 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 article’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* article by F&H (2004) does not provide a response and, hence, that any speculation based on this article (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 F&H’s study, is it still worthwhile to pursue experimental investigations about 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 (2004) 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 and Chomsky (1963) themselves

noted about a very similar example (see also Blaubergs & Braine, 1974). The literature on self-embedding is devoted entirely 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 in which artificial languages have been used do not authorize greater optimism. The experiment reported above did not reveal any learning, even though relationships were between specific tokens, instead of between 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 non-adjacent dependencies, which is a prerequisite for detecting a center-embedded structure, is possible only 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, Richmond, & Chater, in press; Perruchet, Tyler, Galland, & Peereman, 2004) suggest that learning nonadjacent dependencies occurs only in conditions in which 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 cannot be described by FSGs. Recent works in linguistics, however, tend to reverse this view and to conclude 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 (2004) were right in claiming that humans easily master a PSG, whereas tamarins can only master an 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 an FSG, which is easily mastered by tamarins, according to F&H. Clearly, tamarins do not fail to master only 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 (2004) 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 (1957) can be of much help in 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*. The Hague: Mouton.
- CHRISTIANSEN, M. H., & CHATER, N. (1999). Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science*, **23**, 157-205.
- CLEEREMANS, A. (1993). *Mechanisms of implicit learning: A connectionist model of sequence processing*. Cambridge, MA: MIT Press, Bradford Books.
- 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 R. Alterman & D. Kirsh (Eds.), *Proceedings of the 25th Annual Conference of the Cognitive Science Society* (pp. 270-275). Mahwah, NJ: 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 noncommercial purposes. In H. T. Bunell & W. Isardi (Eds.), *Proceedings of the Fourth International Conference on Spoken Language Processing* (pp. 1393-1396). Wilmington, DE: Dupont Institute.
- 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, J. (1999). Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language & 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. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 2, pp. 419-493). New York: Wiley.
- 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.
- MORGAN, J. L., & NEWPORT, E. L. (1981). The role of constituent structure in the induction of an artificial language. *Journal of Verbal Learning & Verbal Behavior*, **20**, 67-85.
- 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., RICHMOND, K., & CHATER, N. (in press). Phonology impacts segmentation in online speech processing. *Journal of Memory & Language*.
- 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). Cambridge, MA: MIT Press.
- PERRUCHET, P., TYLER, M. D., GALLAND, N., & PEEREMAN, R. (2004). Learning nonadjacent dependencies: No need for algebraic-like computations. *Journal of Experimental Psychology: General*, **133**, 573-583.
- SANTELMANN, L. M., & JUSCZYK, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition*, **69**, 105-134.

NOTES

1. In fact, Chomsky (1957) included among the productions that could not be described by an FSG 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 the latter from the center-embedded structures, which F&H (2004) designated as their main target throughout their article, 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 $n = 3$, **AAABBB**, instead of **ABABAB** (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 (2004) one. However, this possibility was not detrimental for our objective, and in fact, to anticipate, there was no sizable difference between the performance of our and F&H's subjects on this aspect.

3. Mark Liberman has proposed an alternative hypothesis (see <http://itre.cis.upenn.edu/~myl/language/og/archives/000355.html>).

4. Commenting on their own results with humans, F&H (2004) noted, "these data are consistent with other experimental findings that humans can learn a PSG and appear to prefer phrase-structured input" (p. 379). Now, the empirical studies to which they referred to support their claim (Morgan, Meier, & Newport, 1989; Morgan & Newport, 1981) used an artificial language generated by an FSG, and not by a PSG (Morgan et al., 1989, indeed used the expression *phrase-structured input*, but in a sense unconnected to Chomsky's, 1957, *phrase structure grammar*). Moreover, the aim of these articles was to show 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 sentence. 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 FSG approximation (e.g., Pereira & Wright, 1997) or simple recurrent networks (Christiansen & Chater, 1999).

APPENDIX

| Study Sentences | | Test Sentences | |
|-------------------------|---------------------------|-----------------|-----|
| 1 <u>ba</u> lanosadogu | 17 vutu <u>ba</u> gulinu | 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 vumib <u>ba</u> gumonu | 7 miladidomoka | (I) |
| 8 novumimonusa | 24 <u>ba</u> vuladonugu | 8 robabigu | (I) |
| 9 <u>ba</u> divunukagu | 25 mib <u>ba</u> nosagumo | 9 baturobiligu | (C) |
| 10 mit <u>ba</u> gulimo | 26 milanosadomo | 10 norotubisali | (I) |
| 11 vuromimobinu | 27 tu <u>ba</u> guli | 11 dinosaka | (C) |
| 12 divunuka | 28 tunosali | 12 rovonubi | (C) |
| 13 tuladoli | 29 ditulika | 13 dimikamo | (I) |
| 14 lavunudo | 30 vub <u>ba</u> robigunu | 14 nolatulidosa | (C) |
| 15 <u>ba</u> novunusagu | 31 rob <u>ba</u> gubi | 15 lavudonu | (I) |
| 16 ladikado | 32 dirobika | 16 vutulalinudo | (I) |

Note—The table displays the study and the test sentences heard by Subject 1. For this subject, *ba* was consistently paired with *gu* in such a way that if *ba* occurred as A3, *gu* occurred as B3 (e.g., Study Sentence 1), if *ba* occurred as A2, *gu* occurred as B2 (e.g., Study Sentence 27), and if *ba* occurred as A1, *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 with *gu* (more generally, the second half of this sentence should have been *bikagu*, instead of *kagubi*).

(Manuscript received March 23, 2004;
revision accepted for publication July 5, 2004.)