Muriqui monkey vocalizations are complex: formal and statistical assessment

Keny Chatain[a,1], **Philippe Schlenker**[b,c,†], **Didier Demolin**[d,e,f], **Francisco D.C. Mendes**[g], **Robin J. Ryder**[8,†], **and Emmanuel Chemla**[9,†]

[a]Department of Linguistics and Philosophy, MIT.; [b]Institut Jean-Nicod, Département d'Etudes Cognitives, ENS, PSL University, EHESS, CNRS.; [c]Department of Linguistics, New York University, New York, New York.; [d]Université Sorbonne Nouvelle, Paris 3; [e]ILPGA (Institut de linguistique et de phonétique générale et appliquée); [f]Laboratoire de Phonétique et Phonologie (LPP), CNRS UMR 7018; [g]Departamento de Processos Psicológicos Básicos-IP, Universidade de Brasília, Brasília, Brazil; Av. Universitária 1440, Cx. Postal 86, 74605-010 Goiânia, GO, Brasil; [h] Centre de Recherche en Mathématiques de la Décision, CNRS, UMR 7534, Université Paris-Dauphine, PSL University, Paris, France. ; [i] Laboratoire de Sciences Cognitives et Psycholinguistique, Département d'Etudes Cognitives, ENS, PSL University, EHESS, CNRS.

This manuscript was compiled on June 2, 2020

Recent recordings of Muriqui monkeys display numerous "increasing sequences" with increasing number of p calls interspersed with z calls (e.g. *zp-zpp-zppp-zpppp...*), with *p* and *z* two types of call units. We prove that this class of sequences is outside the current computational bound postulated for human languages (*i.e.* mildly context-sensitive languages). We then develop a novel statistical analysis of the entire corpus of Muriqui vocalizations and establish that their language is not "finite-state", a class of languages within which animal languages are usually thought to lie. This result warrants a reconsideration of accepted views on the limitations of animal syntax.

Animal communication | Human language | Chomsky's hierarchy | Formal language theory

A classical tool to measure the complexity of a language is the Chomsky hierarchy (1) illustrated in Fig. 2. Intuitively, the place of a language in this hierarchy characterizes the minimal computational resources necessary to identify its acceptable sentences: is some form of memory required to track or count specific elements? Does the language impose dependencies between segments far apart in a sentence? The proposed formal approach applies to English, other human languages or any type of system that can be formally described as a language. We argue that the language of the Muriqui monkeys (Brazil) lies very high in the Chomsky hierarchy of complexity of languages.

An epoch-making result was that English goes beyond the "finite-state" bound (1). Intuitively, a finite-state machine parses a sentence by visiting successive "states" (Fig. 2): the machine is initialized in a starting state, and moves to its next state using any path labeled with the next word to be parsed. The state reached after parsing the final word in a sentence indicates whether the whole sentence is acceptable or not. The result is that no such machine can capture English syntax. This is based on the application of a so-called "pumping lemma", a version of which we will apply to Muriqui vocalizations. Intuitively, the idea is that if a finite-state machine accepts a sentence S with more than n elements, then either the machine has more than n states, or some state is visited several times during parsing of S. If the latter, then there is a loop L in the parsing path of S, and this entails that the machine also accepts some longer sentences, those with a similar parsing paths where the loop L is traversed more than once (cf. Fig. If a language does not contain sequences formed from sequences longer than n formed in that manner, then it cannot be described by a finite state machine. Such an argument has been applied to prove that English is not finite-state (2, 3).

Applying a similar type of argument to the inferred description of the Muriquis, we will show that the Muriqui language list at a level of the hierarchy above what has been formally proven for human languages (4). Still, such arguments have an intrinsic weakness: no set of observed data is infinite, and every finite set can be produced by a finite-state machine (5, 6). In the case of human language, one typically argues that despite such limitations, speakers' knowledge is most insightfully described as rules (such as relative clause formation or, more concretely, embedding under "John thinks that...", cf. Fig. 2). Through rules, we can determine whether an infinite number of sentences are part of 'idealized' English, even if they are arbitrarily long, and we can thus establish that finite-state machines cannot capture these non-finite languages.

The latter argument may be applied to animals by asking what rules they can master in learning tasks (7–12), but it is an indirect argument about their actual communication system. For communications systems, alternative statistical methods have been deployed to prove that some communication systems were beyond the finite-state bound (13). In this work, we develop a stronger case by first exhibiting a specific pattern in the language that is provably of high complexity (even

Significance Statement

We show that particular call sequences of Muriqui monkeys have a spectacularly high degree of complexity, as measured by Chomsky's hierarchy of formal languages. We present formal arguments that this pattern is beyond the level standardly assumed for human languages. Taking advantage of this particular pattern, we are able to conduct a direct statistical test on our corpus, which proves that it cannot be effectively recognized by a finite-state machine. This statistical approach avoids difficulties with extracting formal description from finite corpora. Taken together, these results challenge attempts to characterize the uniqueness of human languages in terms of Chomsky's hierarchy only.

Didier Demolin and Francisco D. Mendes collected the corpus of Muriqui vocalizations and performed the phonetic analysis. Keny Chatain proposed the formal proof of context-sensitiveness, which was checked for correctness by Emmanuel Chemla and Robin Ryder. Keny Chatain wrote the code for the statistical analysis, under the supervision of Emmanuel Chemla and Robin Ryder. Keny Chatain, Emmanuel Chemla, Philippe Schlenker and Robin Ryder wrote the manuscript and incorporated remarks, additions and corrections suggested by Didier Demolin and Francisco D. Mendes.

There are no competing interests.

† Authors two, five and six are joint senior co-authors.

[1]To whom correspondence should be adressed. E-mail: kchatain@mit.edu

higher than the finite state bound). Second, exhibiting these sequences allow us to construct unique statistical measures, which can be used to locate the whole language within the Chomsky hierarchy.

Data

It was claimed (14) that the call sequences of Muriqui monkeys can only be produced with context-sensitive grammars; we evaluate this claim theoretically and statistically. 647 sequences were recorded and transcribed by experts, who found that 14 units (elementary sounds akin to phonemes or words) recurred across sequences, including three units labeled r, h and p. These play a central role because the corpus contains a large number of sequences of the form *hprpphppp*, with gradually increasing numbers of p's prefixed with lone r's or h's (cf. Fig. 4). We use z to refer to a unit that is either h or r. If we call 'chunk' any maximal zp^+ string (where p^+ represents an arbitrary number of p's), these chunks are 'increasing' in their number of p's. Formally, we say a sequence is increasing if each chunk is the same length or one p longer than the previous chunk, i.e. if it is of the form

$$w_1(zp)^{a_1}(zp^2)^{a_2}\dots(zp^n)^{a_n}w_2 \text{ with } n \geq 1 \text{ and } a_1,\dots,a_n \geq 1.$$

(where w_1 and w_2 are any sequences that do not contain r, h or p).

Results

What is the theoretical import of increasing sequences? We will (i) prove that the sub-language they define goes beyond the finite-state bound (and in fact requires context-sensitive grammars), and then (ii) provide a statistical argument that the observed frequency of such increasing sequences makes it unlikely that the entire language is finite-state: Muriquis genuinely require that we rethink accepted limits on animal syntax.

The language of increasing sequences is not finite-state, and not even mildly context-sensitive. While the language of increasing sequences can be produced by a context-sensitive grammar, it cannot be produced by finite-state grammars. It cannot even be produced by a mildly context-sensitive grammar, although this more powerful formalism is thought to be sufficient to handle human syntax (cf. Fig. 1, (4)). We prove this in two steps (details provided in Supporting Information appendix):

First, we show that the language is at most context-sensitive by exhibiting a 'Linear Bounded Automaton' (LBA) that recognizes the target language. LBAs are Turing machines with space limitations; they are more general than finite-state machines (15), and they recognize exactly the context-sensitive languages. Like standard Turing machines, LBAs have a tape (i.e. a memory) and a set of instructions that allow them to read an input one symbol at a time, rewriting a symbol and moving left or right. The additional constraint for an LBA is that for some integer K, it only uses a tape which is K times the size of the input – the required memory is thus linear in the input size, whence the name. Recognizing an increasing sequence doesn't require much memory: the LBA scans the sequence from left to right, one zp^+ chunk at a time, stores the latest sequence of p's in memory, checks that the next chunk contains zero or one more p,and rejects the sequence otherwise. This new p^+ replaces the previous one in memory, and the process is iterated until the end of the sequence, with a memory space that remains shorter than twice the size of the input (so here, $K = 2$).

Second, we show the negative result that the language of increasing sequences (henceforth L_{IS}) is not finite-state. This result is striking because cases of animal syntax of that complexity are rarely found, if at all. The proof of this result is given in the Supplementary Information appendix, but an intuitive version of it can be derived from the 'pumping' property. Suppose L_{IS} is recognized by a finite-state machine with n states. Take a string S in the language with contiguous chunks $zp^{n+1}zp^{n+2}$. Since the machine only has n states, it must go through the same state twice when processing p^{n+1}. This guarantees the existence of a loop (in the sense above), which can be repeated any number of times without changing the acceptability of the resulting string. However, these iterations will transform the first zp^{n+1} chunk into one with more p's than the following zp^{n+2} chunk, which takes us out of the target language. This is in violation of the pumping lemma applicable to finite-state languages.

A stronger negative result is that the language L_{IS} of increasing sequences cannot even be produced by a mildly context-sensitive grammar, although such a formalism is thought to be sufficient to handle human languages (4). This result is formally proved in the Supporting Information appendix using a (more sophisticated version of) the pumping lemma.

The Muriqui language is unlikely to be finite-state. We have shown that the sub-language of increasing sequences is context-sensitive, but this argument alone is not sufficient to show that the full Muriqui language is not finite-state. Indeed, a sub-language can be of greater complexity than the entire language (16). For example, one might worry that the Muriqui language contains $(zp^+)^+$, the language of sequences obtained by concatenation of $zp\dots p$ sequences, and increasing sequences are an 'accidental' subpart of that language.

We address this worry by adopting a probabilistic approach. Moving from finite-state grammars to probabilistic finite-state grammars (with each transition being assigned a probability), as in classic Markov processes, we can ask not just whether a finite-state machine could *in principle* produce the increasing sequences we find in the corpus (along with many other patterns), but whether it could produce them *with the frequency that we find*. Our findings are summarized in Fig. 5, detailed in the Supporting Information appendix, and we explain them now in prose.

A probabilistic version of the finiteness objection still holds: since our corpus is finite, it is always possible to find a finite-state probabilistic grammar that generates increasing sequences with frequencies matching those found in the corpus. However, this grammar may have an implausibly large number of states. To quantify this "implausibility", we make use of a classic penalty known as the Bayesian Information Criterion (BIC), which provides a principled trade-off between goodness of fit to the sample and parsimony of parameters (here, roughly, the number of states squared).

Using this criterion, we find the finite-state grammar which provides the optimal trade-off, simulate some artificial sequences from it, and compare these sequences to our corpus. If the Muriqui grammar were finite-state, we would expect the proportions of increasing sequences in the observed and simulated data to be similar. However, we found that the trained grammar generates significantly fewer increasing sequences (empirical p-value $< 10^{-7}$), which strongly suggests that the Muriqui grammar is not finite-state. In other words, finite-state models systematically underestimate the proportion of increasing sequences. As a control, we checked that we did not find a similar result for otherwise comparable decreasing sequences: such sequences occur at similar frequencies in the observed and simulated data.

The under-estimation of increasing sequences by finite-state models persists even when we impose that the proportion of sequences $(zp^+)^+$ generated by the model be the same as in the corpus; thus we conclude that the under-representation of increasing sequences cannot be blamed on under-representation of the more general regular pattern $(zp^+)^+$.

The discussion above neglects the fact that the sequences in the corpus are noisy. This may result in under-counting increasing sequences in the corpus and in the simulated data. We take this into account by considering that sequences could result from some amount of errors, be they performance or coding errors. The result is robust to that type of noise and still holds if we allow sequences to depart from the actual sequences by up to 20%.

Discussion

Muriqui call sequences appear to be very high in the Chomsky hierarchy: on statistical grounds, the sample of vocalizations is unlikely to have been extracted from a language within the finite state bound, against accepted opinions on the limits of animal syntax (4), and stronger results could be obtained once appropriate statistical methods will be available. Still, it would be hasty to conclude that Muriquis have cognitive abilities that go beyond human syntax. It could be that their cognitive resources just allow them to produce sequences mixing p's and z's without constraint, while some other mechanism (for instance, of a physiological nature) adds constraints of its own.

For instance, it could be that excitement levels vary in systematic ways during the utterance of such sequences, suppose excitement increases, and that this variation is reflected in the typical shape of the sequences, suppose that more excitement leads to more p's. Details matter, however, and the current data do not make such an hypothesis very plausible: excitement would have to affect the p selectively, and we note that the spacing of the z's in the sequences is constant, which is not what is to be expected if excitement levels perturb the strings as they are produced (see the temporal regularity of the p's in Figure 3). But this type of explanation highlights an essential question: it is one thing to place Muriqui calls in a certain part of the Chomsky hierarchy, and quite another to explain why (i.e. through which cognitive or non-cognitive mechanisms) it occupies this particular position.

The Muriqui vocalizations provide a unique case study in the animal kingdom. There we can exhibit a specific pattern of very high complexity in the Chomsky hierarchy (beyond the limits attributed to human languages). Such a discovery could motivate exploration of other measures of complexity that would better characterize the exceptionality of human language.

ACKNOWLEDGMENTS. This research received partial funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 788077, Orisem, PI: Schlenker). Research was conducted at Institut d'Etudes Cognitives, Ecole Normale Supérieure - PSL Research University. The research leading to these results was supported by ANR-17-EURE-0017.We thank César Ades, Karen Strier and Charles Snowdon for their support and advisorship during Mendes' research on muriqui vocal repertoires.

1. N Chomsky, Three models for the description of language. *IRE Transactions on information theory* **2**, 113–124 (1956).
2. N Chomsky, Aspects of the Theory of Syntax, Technical report (1964).
3. B Partee, A ter Meulen, R Wall, *Mathematical methods in linguistics.* (Springer Science & Business Media) Vol. 30, (2012).
4. AK Joshi, How much context-sensitivity is required to provide reasonable structural descriptions: Tree adjoining grammars in *Natural Language Parsing*, eds. DR Dowty, L Karttunen, AM Zwicky. (Cambridge University Press), pp. 206–250 (1985).
5. S DeDeo, Collective phenomena and non-finite state computation in a human social system. *PloS one* **8**, e75818 (2013).
6. C De la Higuera, *Grammatical inference: learning automata and grammars.* (Cambridge University Press), (2010).
7. WT Fitch, MD Hauser, Computational Constraints on Syntactic Processing in a Nonhuman Primate. *Science* **303**, 377 LP – 380 (2004).
8. P Perruchet, A Rey, Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates? *Psychon. Bull. & Rev.* **12**, 307–313 (2005).
9. TQ Gentner, KM Fenn, D Margoliash, HC Nusbaum, Recursive syntactic pattern learning by songbirds. *Nature* **440**, 1204 (2006).
10. C Van Heijningen, J De Visser, W Zuidema, C Ten Cate, Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proc. Natl. Acad. Sci.* **106**, 20538–20543 (2009).
11. K Abe, D Watanabe, Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nat. neuroscience* **14**, 1067–1074 (2011).
12. GJL Beckers, JJ Bolhuis, K Okanoya, RC Berwick, Birdsong neurolinguistics: Songbird context-free grammar claim is premature. *Neuroreport* **23**, 139–145 (2012).
13. A Kershenbaum, et al., Animal vocal sequences: not the Markov chains we thought they were. *Proc. Royal Soc. B: Biol. Sci.* **281**, 20141370 (2014).
14. D Demolin, C Adès, FD Mendes, Non linear phenomena in muriqui vocalizations in *18th International Congress on Sound and Vibration (ICSV 18).* pp. pp 9–19 (2011).
15. JE Hopcroft, *Introduction to automata theory, languages, and computation.* (Pearson Education India), (2008).
16. GK Pullum, Such That Clauses and the Context-Freeness of English. *Linguist. Inq.* **16**, 291–298 (1985).

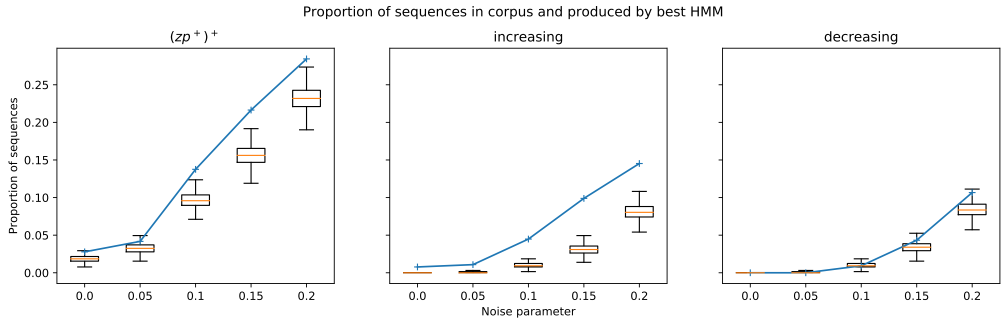

Fig. 1. Frequencies of $(zp^+)^+$ sequences, increasing sequences and decreasing sequences in the Muriqui corpus at different levels of noise (plain line). Corresponding numbers are given for samples obtained from fitted probabilistic finite-state grammars (box and whisker plots). Increasing sequences are specifically underestimated by these finite-state grammars (middle plot).

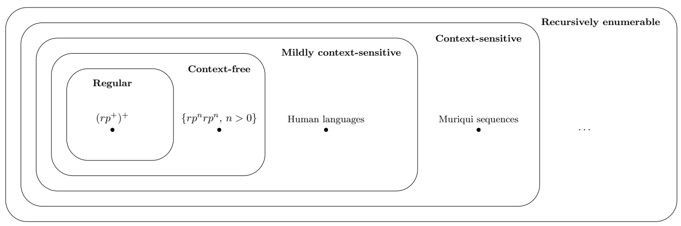

Fig. 2. Chomsky hierarchy with examples of languages from the different classes, including the current results about Muriqui sequences.

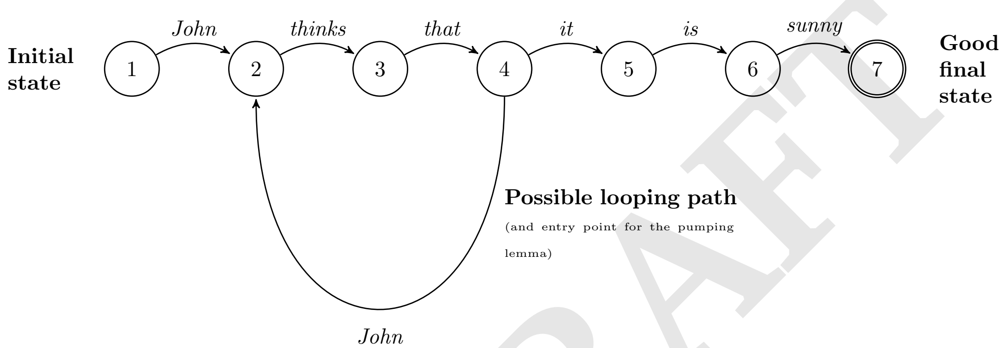

Fig. 3. A finite-state machine can parse sentences such as "John thinks that John thinks that (...) it is sunny", illustrating a possible application of the pumping lemma.

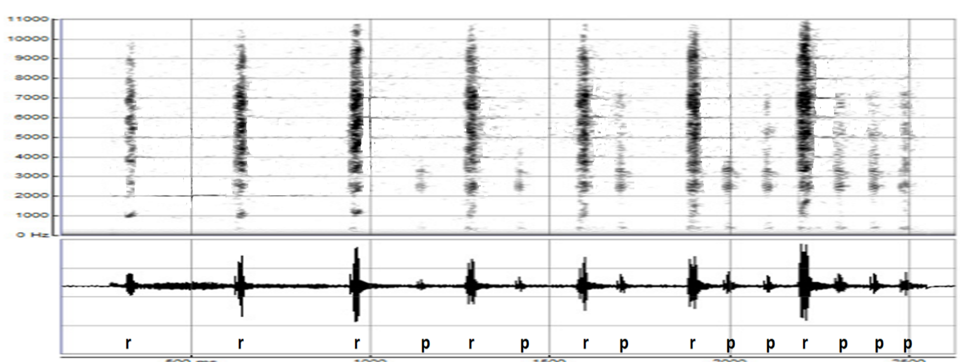

Fig. 4. Spectrogram of an increasing sequence from the Muriqui corpus.

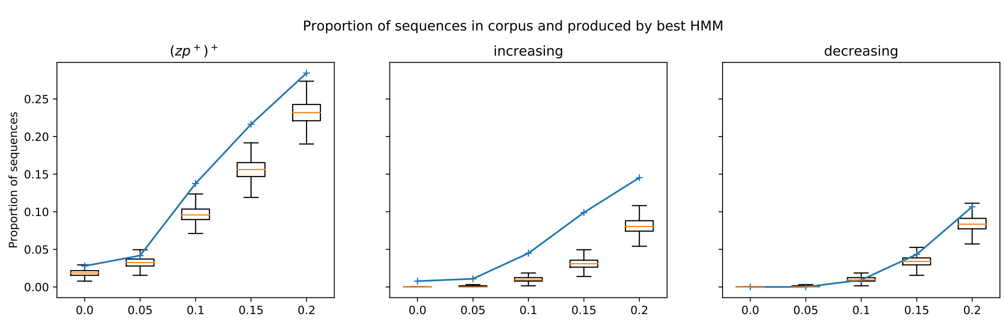

Fig. 5. Frequencies of $(zp^+)^+$ sequences, increasing sequences and decreasing sequences in the Muriqui corpus at different levels of noise (plain line). Corresponding numbers are given for samples obtained from fitted probabilistic finite-state grammars (box and whisker plots). Increasing sequences are specifically underestimated by these finite-state grammars (middle plot).