Evidence of an evolutionary precursor to human language affixation in a non-human primate

Ansgar D. Endress*, Donal Cahill, Stefanie Block, Jeffrey Watumull and Marc D. Hauser

Department of Psychology, Harvard University, Cambridge, MA 02138, USA
*Author for correspondence (ansgar.endress@m4x.org).

Human language, and grammatical competence in particular, relies on a set of computational operations that, in its entirety, is not observed in other animals. Such uniqueness leaves open the possibility that components of our linguistic competence are shared with other animals, having evolved for non-linguistic functions. Here, we explore this problem from a comparative perspective, asking whether cotton-top tamarin monkeys (Saguinus oedipus) can spontaneously (no training) acquire an affixation rule that shares important properties with our inflectional morphology (e.g. the rule that adds -ed to create the past tense, as in the transformation of walk into walk-ed). Using playback experiments, we show that tamarins discriminate between bisyllabic items that start with a specific 'prefix' syllable and those that end with the same syllable as a 'suffix'. These results suggest that some of the computational mechanisms subserving affixation in a diversity of languages are shared with other animals, relying on basic perceptual or memory primitives that evolved for non-linguistic functions.

Keywords: Animal cognition; evolution of language; morphology; language acquisition

1. INTRODUCTION
While it is clear that only humans have a language faculty, it is less clear which components of this system are unique to humans, and which unique to language. In fact, although attempts to teach non-human animals to produce simplified languages have largely failed (Terrace et al. 1979; Savage-Rumbaugh et al. 1993), and studies of their natural communication show only weak evidence of homologous or analogous abilities (Hauser 1996; Liebal et al. 2004; Cheney & Seyfarth 2005; Arnold & Zuberbühler 2006; Suzuki et al. 2006), different animals show perceptual competences that may well feed into language processing in humans (Kuhl & Miller 1975; Kluender et al. 1987; Ramus et al. 2000).

Here, we build on the above tradition exploring aspects of perceptual competence, asking whether animals have non-linguistic abilities that are necessary for some forms of language-specific, grammatical computations (Hauser et al. 2001; Fitch & Hauser 2004; Gentner et al. 2006; Murphy et al. 2008). We start from the observation that, across the world's languages, morphological transformations adding verbal material to the word-edges (i.e. prefixation and suffixation) are much more frequent than transformations adding verbal material in other positions (Greenberg 1957). For example, the English past participle is formed by adding the 'ed' suffix to the end of a stem (as in talk-ed), while the German past participle is formed by adding the 'ge' prefix to the beginning of a stem and either the 'en' or the 't' suffix to its end (as in ge-sag-t, 'said'). In these and other languages, word-edges appear well suited for some linguistic transformations (Nespor & Vogel 1986; McCarthy & Prince 1993).

Here we ask whether a non-human animal—the cotton-top tamarin monkey—has the requisite mechanisms for learning formally similar prefixation and suffixation patterns. Our goal, therefore, is not to show that animals such as tamarins have language, but rather, that certain components of our expressed languages rely on domain-general mechanisms of learning and memory that are likely to be shared with other animals, including, we suggest, the capacity to extract patterns of temporal ordering.

In brief, we exposed subjects to a sequence of bisyllabic items conforming to a common pattern. For example, they heard a sequence of 'stem' syllables all preceded by the same prefix syllable. Following this familiarization, they were exposed to new bisyllabic items. Half were preceded by the same prefix syllable as during familiarization, and half were followed by that syllable, and thus violated the familiarization pattern. We asked whether tamarins would respond more to bisyllabic items violating the familiarization pattern than to items consistent with it.

2. MATERIAL AND METHODS

The detailed methods are described in Hauser et al. (2001); here, we highlight only critical differences.

(a) Participants
We tested 14 adult tamarins (seven males; mean age 8.2 years) socially housed in a colony room. For medical reasons, one subject completed only the suffixation condition, and one only the prefixation condition.

(b) Materials
We used naturally recorded syllables as stimuli from native speakers of American English. The affix syllable was always 'shoy' uttered by a male speaker. The familiarization stems (see below) were 'bi, ka, na, to, gu, lo, ri and nu', pronounced by a female speaker, and 'ba, pu, di, ki, lu, ro and mo' pronounced by a male speaker with a lower voice than that of the speaker of the affix syllable. We used a mixture of different speakers of different genders to prevent subjects from using low-level cues (such as pitch differences between vowels) for their generalizations.

The test stems were the syllables 'brain, breast, wisp, snake and swan', all pronounced by a different female speaker; we used words because speakers found it easier to read English words than phonemic transcriptions.

Syllables were recorded individually, normalized to a duration of 400 ms and then RMS amplitude normalized.

(c) Design
We first familiarized subjects to bisyllabic items conforming to either a prefixation or suffixation pattern, and then tested them on new items that either violated or were consistent with the familiarization pattern. Our dependent measure was an orienting response (see below) towards the speaker playing back a test item. Based on prior work using the same method, we predicted that tamarins would orient more to violations of the familiarization pattern than to items consistent with it.
Half of the subjects were first tested with the prefixation pattern, and 29 days later with the suffixation pattern. The other half was first tested on the suffixation pattern, and 33 days later with the prefixation pattern.

(d) Familiarization
During the familiarization phase, subjects heard a sequence of bisyllabic items (hereafter ‘words’) that all conformed to a common pattern. In the prefixation condition, all words were composed of the prefix ‘shoy’ and one of the familiarization stems mentioned above (e.g. ‘shoy-bi’, ‘shoy-mo’). In the suffixation condition, all words were composed of a familiarization stem and the suffix ‘shoy’ (e.g. ‘bi-shoy’, ‘mo-shoy’). There was no silence between the prefix and the stem, and words were separated by silences of 2 s.

The evening before being tested, monkeys not participating in a condition were brought out of the colony room. Then, the familiarization stream was played to the remaining monkeys through speakers inside the colony room.

The 14 words were played 70 times, yielding a familiarization duration of 29.4 min. Words followed each other in random order with no repetitions.

(e) Test
The morning following this familiarization, subjects were transferred from their home cage to a test cage inside a sound-attenuated chamber. Before proceeding to the test phase, they were given a refresh familiarization of 2.1 min consisting of five repetitions of the 14 familiarization words.

During test, subjects typically clung to the wire mesh on the front of the test cage, facing the camera. Stimuli were played through a concealed speaker. Stimuli consisted of the five test stems mentioned above. Each stem was presented twice, once with the prefix ‘shoy’, and once with this syllable as the suffix. Stimuli were arranged in a list alternating prefixed and suffixed stems. Half of the subjects started with a prefixed stem, and half with a suffixed stem.

(f) Coding
We counted the orientation responses to stimuli consistent with or violating the familiarization pattern. Orientation responses were counted if, within a 2.8 s window following the stimulus onset (corresponding to a 2 s window following the stimulus offset), the subject performed a head rotation of at least 60° in the horizontal plane, and if the subject’s looking direction was not below that plane at the end of the rotation. Trials were excluded if the subject looked in the direction of the speaker at the start of the trial, jumped during the 2.8 s period from the trial onset, or vocalized during the stimulus.

Trials were started at least 10 s and at most 60 s after the beginning of the previous trial by an experimenter blind to the trial type (consistent or violation). Trials were started when the subject looked in the direction of the speaker at the start of the trial, jumped during the 2.8 s period from the trial onset, or vocalized during the trial.

All sessions were coded offline, independently and blindly by three experimenters. Average inter-observer agreement was 79.6 per cent, Cohen’s $k = 0.68$. (Cohen’s $k$ is a measure of inter-observer agreement that takes into account the frequency of agreements that occur by chance.)

To reach a complete consensus, we reviewed all trials for which there was no uniform agreement until all experimenters could agree on the response measure. If no consensus could be reached, the corresponding trial was removed from analysis ($n = 2$ out of 260 trials). We believe that the final consensus is much more reliable than judgments of individual experimenters; indeed, if a coder misses a criterion with probability $p$, all three coders miss it with probability $p^3$.

To assess the reliability of the final consensus, an experienced observer who had not previously coded the data (M.D.H.) coded a randomly selected set of 25 trials. He agreed with the consensus codes in 92 per cent of the trials (Cohen’s $k = 0.872$).

3. RESULTS
For each monkey, we computed the proportion of orienting responses to violations of the familiarization pattern and to test items consistent with that pattern, respectively. For the monkeys completing both conditions ($n = 12$), we submitted these proportions to an analysis of variance (ANOVA) with the within-subject factors, item-type (consistent versus violation) and condition, (prefixation versus suffixation) and the between-subject factor, condition order. This ANOVA yielded a significant main effect of item-type ($F_{(1,10)} = 7.43$, $p = 0.021$, $\eta^2 = 0.413$), but no other main effects or interactions (all $p$’s $> 0.05$). We thus pooled the proportions from all conditions and all subjects.

Overall, monkeys (including those participating in only one condition) oriented significantly more to violations (proportion of orientations: $M = 0.519$, s.d. $= 0.192$) than to consistent items ($M = 0.370$, s.d. $= 0.253$), $F_{(1,13)} = 5.07$, $p = 0.042$, $\eta^2 = 0.280$ (repeated-measures ANOVA). Of the monkeys responding more to either consistent items or violations, nine out of 11 oriented more to violations ($p = 0.033$, one-tailed binomial test) (figure 1).

4. DISCUSSION
Our results suggest that, in the absence of training, cotton-top tamarins learn a rule that is formally similar to affixation patterns (i.e. prefixes and suffixes) in natural language. These results cannot be explained by a simple association for two reasons. First, because the stems used during test were maximally dissimilar from those used during familiarization, subjects must have generalized the affixation rule to new stems, as opposed to recalling the position of particular stems. Second, it is highly unlikely that subjects associated the test stems with the affix. As the monkeys had never heard the test stems together with the affix, they could not have associated the test stems with the affix through prior exposure.

Given that both humans and cotton-top tamarins can learn this particular aspect of affixation patterns,
one may ask how each species computes these patterns. We suggest that the most plausible account refers to the psychological mechanisms that are used to process affixation patterns, and specifically, mechanisms that are shared in different domains across human and non-human species. When humans learn such forms during language acquisition, however, they must link these domain-general mechanisms of learning and memory to our distinctively linguistic phonological, syntactic and semantic processes and representations; in contrast, no other animal can link these forms to such representations and processes. In linguistic terms, non-human animals may have the capacity to learn surface transformations involved in affixation, but they cannot link them to other aspects of linguistic structure. We conclude by making a few brief remarks on this general thesis.

As noted in §1, morphological affixation patterns tend to place verbal material either at the beginning or the end of words, and thus at the word-edges (Greenberg 1957). From a computational perspective, however, edges are just the sequential positions that can be encoded particularly well (since all positions are encoded relative to the sequence-edges, see Henson (1998)), a conclusion that seems to hold for other primates, including chimpanzees (Endress et al. submitted) and potentially rhesus monkeys (Orlov et al. 2000; Terrace et al. 2003). Hence, in line with previous proposals (Endress et al. 2005, in press), we suggest that the language faculty uses similar positional mechanisms to compute affixation patterns, and although these mechanisms are uniquely used in humans to create and understand words, the mechanisms themselves are not specific to humans or language. For example, when infants acquire the morphological distinction for marking the past tense, they may simply recognize, like other primates, that this distinction entails placing the ‘ed’ morpheme in the right edge of words, although they (and other animals) can use similar positional mechanisms in a variety of non-linguistic domains. Unlike other primates, however, infants can use such evolutionarily ancient abilities for purposes that are specifically linguistic and (presumably) unique to humans.

All research was approved by the Animal Care and Use Committee at Harvard University (protocol number 92-16).

A.D.E. and M.D.H designed the research, analysed data, and wrote the paper; A.D.E., D.C., S.B. and J.W. performed research. Funding for this work was provided by M.B.B. grants to M.D.H. and A.D.E. as well as gifts from J. Epstein and S. Shuman and a McDonnell foundation grant to M.D.H. We thank A. Caramazza and S. Pinker for helpful comments on earlier versions of this manuscript.