Primate vocal behavior is often considered irrelevant in modeling human language evolution, mainly because of the caller’s limited vocal control and apparent lack of intentional signaling. Here, we present the results of a long-term study on Campbell’s monkeys, which has revealed an unrivaled degree of vocal complexity. Adult males produced six different loud call types, which they combined into various sequences in highly context-specific ways. We found stereotyped sequences that were strongly associated with cohesion and travel, falling trees, neighboring groups, nonpredatory animals, unspecified predatory threat, and specific predator classes. Within the responses to predators, we found that crowned eagles triggered four and leopards three different sequences, depending on how the caller learned about their presence. Callers followed a number of principles when concatenating sequences, such as nonrandom transition probabilities of call types, addition of specific calls into an existing sequence to form a different one, or recombination of two sequences to form a third one. We conclude that these primates have overcome some of the constraints of limited vocal control by combinatorial organization. As the different sequences were so tightly linked to specific external events, the Campbell's monkey call system may be the most complex example of 'proto-syntax' in animal communication known to date.

One way of studying language evolution is to compare the communicative abilities of humans and animals. Parallels with human language can be found at various levels, both in terms of production and comprehension. Particularly relevant are cases of social influences on vocal development (1), cases of infant babbling (2), and other types of vocal learning (3–5). In some species, there is evidence for population-wide convergence effects in the form of culturally transmitted dialects [e.g., starlings (6), whales (7), and Japanese macaques (8)]. In terms of pragmatic use, there is good evidence that call production can be influenced by specific audiences (9, 10). In terms of comprehension, primates and possibly many other species are able to assign meaning to different calls if there is a strong relation between a call's acoustic morphology and its eliciting context (11–14). In some species, there is some evidence for hemispheric specialization when processing conspecific calls [e.g., horses (15), Campbell’s monkeys (16), rhesus macaques (17), starlings (18), and sea lions (19)].

Despite all this evidence, there is currently a wide consensus that human language differs from animal communication in a profound way, because the essence of human language is its complex grammatical organization, something that is widely lacking in all animal communication systems (20–22). Despite considerable effort devoted to the topic, the evolutionary origins of this key property of language have remained relatively elusive (23). This is not to say that animal communication does not follow certain combinatorial principles. Gibbons, whales, and songbirds, for example, combine finite and stereotyped sound elements to form more complicated structures (1, 24–27). In some cases, these structures possess hierarchical organization, although very little is known about the relationship between acoustic structure and communicative function. A typical finding is that if the structure of a sequence is artificially altered, for example by changing the composition or order of elements, the signal tends to lose its communicative function (28–30). Another relevant point is that nonhuman primates are perfectly capable of discriminating human speech composed in different ways [e.g., tamarins (31)] and of comprehending simplified nonverbal forms of human syntax [e.g., apes (32–35)].

In natural contexts, spontaneous call combinations have also been observed in nonhuman primates, although there are only a small number of examples. Chimpanzees combine some of their calls in nonrandom ways, although the communicative function of these combinations remains to be investigated (36). Bonobos produce five acoustically distinct call types in response to different foods, with a predictable relationship between the caller’s food preference and the relative frequency of the different calls (37). In putty-nosed monkeys, adult males produce two loud calls, “pyows” and “hacks,” in a range of contexts, including predation. However, when combining the two calls in one specific way (i.e., a few pyows followed by a few hacks), males give a supplementary message to their group members to move away from the current location (38, 39). In chickadees, songs contain C and D notes, which encode different information, while the number and acoustic variation of D notes provides supporting information about the size and dangerousness of a predator (40, 41). During social interactions, animals hear call combinations even more often, especially when individuals exchange vocalizations during social interactions. Playback experiments have shown that baboons recognize individual callers and respond to the combinations of calls produced by them, rather than the individual calls, and extract meaning from them (42, 43).

The vocal behavior of female Campbell’s monkeys is relatively well studied, particularly in captivity, and has revealed considerable socially determined acoustic plasticity (5, 44–47). Adult male Campbell’s monkeys very rarely participate in the various call exchanges that take place within the group, but instead produce a range of loud and low-pitched calls that carry over large distances, much beyond the immediate group distribution (48, 49). These calls are usually given to serious disturbances and in sequences, some of which have been described in earlier research (49–51). For instance, Gautier and Gautier-Hion (48) suggested that hack sequences function as predator signals while “boom-hack” sequences functioned in spatial coordination. Zuberbühler (49) confirmed experimentally that artificially
adding “booms” before a pure hack sequence “cancelled” the predator message. Recently, we have found that the hacks were an acoustically heterogeneous class of vocalizations, which could be categorized into five structurally and contextually different call types (“krak,” “hok,” “krak-oo,” “hok-oo,” and “wak-oo”) (50) (Fig. 1). This subtle acoustic diversity is achieved by frequency transitions and an optional suffixation principle, which increases the males’ call type repertoire considerably.

Here, we studied the alarm calling behavior of free-ranging Campbell’s monkey males in natural and experimental predator situations. We were particularly interested in how males concatenated their repertoire of six call types into call sequences. To this end, we sought to describe the principles governing the observation of sequences in terms of composition and call order and how the different sequences related to external events.

**Results**

**Call Sequences in Nonpredation Contexts.** In nonpredation contexts, we recorded three distinct call sequence types, (i) a pair of “boom” calls (B) given alone, (ii) a pair of boom calls followed by a series of krak-oo (K+), and (iii) a pair of boom calls, followed by a series of K+ calls, with one to several hok-oo (H+) calls interspersed (Table 1 and Fig. 2).

The first sequence type, unaccompanied pairs of boom calls, was not very common, although we managed to record this sequence from seven males (n = 13 events). When producing unaccompanied pairs of boom calls, the common finding was that the male was not in visual contact with the group, usually some 30–50 m away (12 of 13 events). Campbell’s males prefer peripheral positions within the group where they remain in visual contact with at least one group member (average distance between the male and the nearest female: 4.80 ± 0.61 m; n = 10 per male; four habituated males). We observed that, immediately after this boom call sequence, the group often progressed toward the male (8 of 13 events), similar to previous findings in putty-nosed monkeys (39). More detailed observations were possible for the four habituated males. In 10 of 13 instances of unaccompanied pairs of booms, the group traveled for unusually long distances of ~100 m during the next 30 min at a speed significantly higher than normal (~1 m/min; determined by matched samples collected at the same time of day without male calling; Mann–Whitney U test: N1 = N2 = 10, Z = 1.96, P < 0.05).

The second sequence type consisted of a pair of boom calls followed by a K+ sequence (median overall = 10 calls; range: 4–15 calls: n = 53; Fig. 2). This combined sequence, recorded from all four habituated males, was mainly given to falling trees or branches with no other noticeable disturbance (85%). In the remaining cases (15%), the sequence was given in response to fights between other monkey species in the canopy, although this usually led to branches falling as well.

The third sequence also consisted of a pair of boom calls followed by a K+ sequence (overall median = 12 calls; range: 9–16 calls; n = 76), but here the sequence was interspersed with 1–7 H+ calls (Fig. 2). This combined sequence was recorded from all four habituated males and two semihabituated groups and always in response to neighboring Campbell’s monkey groups or single stranger males, suggesting that it functioned in territorial defense. To further test this hypothesis, we investigated whether the location of call production varied in systematic

![Fig. 1](Image)

**Fig. 1.** Spectrograms of the six call types produced by males. Three calls could be described in terms of an acoustically invariable affix following the call stem. +, indicates that call stem is trailed by a suffix, an acoustically invariable “oo” utterance.

<table>
<thead>
<tr>
<th>Sequence composition</th>
<th>M1</th>
<th>M2</th>
<th>M3</th>
<th>M4</th>
<th>M5</th>
<th>M6</th>
<th>M7</th>
<th>M8</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>4</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>—</td>
</tr>
<tr>
<td>B–K+</td>
<td>29</td>
<td>5</td>
<td>13</td>
<td>6</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>B–K+–H+</td>
<td>18</td>
<td>29</td>
<td>22</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>K</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>—</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>K–K+</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>—</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>K+–W+</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>K+–H–W+</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>—</td>
<td>1</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>K+–H+–W+</td>
<td>9</td>
<td>4</td>
<td>8</td>
<td>—</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

M1 to 4, habituated males observed systematically; M5 to 8, semihabituated males observed opportunistically.
ways. We predicted that sequences given in the periphery of a group’s home range contained significantly more \( H^+ \) calls than sequences given in the center, which turned out to be the case (Fig. 3; Fisher’s exact test, \( P < 0.001 \)).

A fourth vocal pattern in nonpredation contexts was the production of single isolated \( K^- \) calls (\( n = 9 \); three habituated males). This always happened when a male was startled by sudden movements of a nonpredatory animal, such as a duiker, flying squirrel, or a human observer.

**Call Sequences in Predation Contexts.** In predation contexts, which is when monkeys interacted with leopards or crowned eagles, we recorded a more diverse range of call sequence types (Table 1 and Fig. 2).

**General Predator Alert.** A first sequence, recorded from all four habituated males and one semihabituated male, consisted of \( K^- \) calls only (median = 15 calls; range: 3–25; \( n = 18 \); Fig. 2). This sequence was rare and given to any auditory sign of a predator, typically after hearing Diana monkey alarm calls (\( n = 15/18 \)) or, more rarely (\( n = 3/18 \)), after hearing predator vocalizations, but never to any visual signs of a predator.

**Crowned Eagle Alarm.** The majority of call sequences to crowned eagles were composed of wak-oo (\( W^+ \)) and krak-oo (\( K^- \)) calls, sometimes with the addition of hok (\( H^- \)) and hok-oo (\( H^+ \)) calls (median = 25 calls; range: 15 to 40; \( n = 38 \); Fig. 2). By analyzing in more details the sequences produced by the three most habituated males, we found that high levels of urgency were associated with a high proportion of \( H^- \) and \( H^+ \) calls in the sequence. There was a significant difference in the proportion of \( H^- \) and \( H^+ \) calls if the caller spotted a real eagle or encountered the eagle model compared to when he only heard eagle vocalizations or Diana monkey eagle alarm calls (Fisher’s exact test, \( P < 0.001 \); Fig. 4).

**Leopard Alarms.** Call sequences to leopards were always composed of krak (\( K^- \)) calls, sometimes combined with krak-oo (\( K^- \)) calls (median = 21 calls; range: 12–35; \( n = 26 \); Fig. 2). In the sequences recorded from the three most habituated males, we found that the level of urgency was associated with a high proportion of \( K^- \) calls in the sequence. Sequences with just \( K^- \) calls were given in response to real leopards and leopard models. \( K^- \) calls were far more common when callers responded to leopard vocalizations or Diana monkey leopard alarm calls. Compared to

---

**Fig. 2.** Composition of call sequences in different behavioral contexts. “Alarm” indicates leopard or eagle alarm calls given by sympatric Diana monkeys.

**Fig. 3.** Geographical distribution of call sequences given during intergroup encounters (A) and after falling of trees or large branches (B). The two study groups were habituated to human observers and had adjacent territories. Each square represents a 100 × 100 m area. We divided each home range into two parts: center (dark squares) and periphery (light squares). The two groups were surrounded by four other semihabituated groups. Stars indicate the location of calling sequences (A, \( n = 71 \); B, \( n = 53 \)).
real leopards and leopard models, K calls were given significantly less often to leopard vocalizations and Diana monkey leopard alarm calls (Fisher’s exact tests, P < 0.001; Fig. 4B).

**The Importance of Call Order.** The different call sequences were not randomly assembled but ordered in specific ways, with entire sequences serving as units to build more complicated sequences. As mentioned, pairs of booms alone instigate group movements toward the calling male, while Kₐ calls functioned as general alarm calls. If combined, the resulting sequence carried an entirely different meaning, by referring to falling wood. In all cases, the booms preceded the Kₐ calls. We also found that another sequence, the H⁻ series, could be added to boom–Kₐ sequences, something that callers did when detecting a neighboring group. H⁻ series were never given by themselves.

Nonpredation events were characterized by the production of two boom calls, which could be given alone (to indicate group movement) or which could introduce subsequent calls (100%, n = 142 cases, all eight males). In response to tree falls, the booms preceded a series of Kₐ calls. In response to neighbors, the booms preceded a series of H⁻ calls, followed by a series of Kₐ calls. In some cases, callers inserted a single Kₐ call between the booms and H⁻ call series. Interestingly, these insertions were only heard in group 1, both by the first male and also by his successor (males 1 and 2), but never in group 2 (i.e., males 3 and 4). Although we only have data from four males, it is possible that an identifying specific feature of group 1. We never heard this particular insertion in the semihabituated groups. Another unexpected difference was that both immigrant males produced significantly more H⁻ calls (males 2 and 4; median = 5; range: 1–7; n = 36) than their predecessors (males 1 and 3; median = 2; range: 1–5; n = 40; Mann–Whitney U tests, Z = 5.53, P < 0.001).

To leopards, males sometimes produced pure Kₐ sequences (n = 9/26), but if Kₐ calls were added, then typically toward the end of the sequence. We compared the number of Kₐ and Kₐ calls in the first and second half of the sequences and found a significant difference (Fisher’s exact test, P < 0.001). To crowned eagles, most sequences began with a series of H (16 of 38 sequences) and typically ended with a series of Kₐ (36 of 38 sequences). This specific order probably reflected the decrease in urgency or threat perceived by the caller. If W⁺ or H⁻ were produced, then they appeared more or less randomly throughout the sequence without any detectable patterns, while H and Kₐ followed the ordering outlined before. There was a significant difference between the number of H and Kₐ at the beginning and at the end of sequences (Fisher’s exact test, P < 0.001; n = 38), which had the effect that the distribution patterns differed significantly between the four call types (χ² test X² = 311.3; DL = 21; P < 0.001).

**Discussion**

In earlier work, we have shown that Campbell’s monkey males use an affixation rule to increase the size of their call repertoire. Adding the suffix “oo” to krak (K → Kₐ) or hok (H → H⁻) calls altered these calls’ meanings in predictable ways (50). Here, we describe regularities at another level, i.e., in how monkeys combined this repertoire of six call types into nine distinct call sequences (Fig. 2). These call combinations were not random, but the product of a number of principles, which governed how semantic content was obtained.

We found five main principles that governed these relationships. First, callers produced sequences composed of calls that already carried narrow meanings (e.g., K = leopard; H = crowned eagle) (49, 50, 52). In these instances, sequence and call meanings were identical. Second, callers produced meaningful sequences, but used calls with unspecific meanings (e.g., Kₐ = predator). Third, callers combined two meaningful sequences into a more complex one with a different meaning (e.g., B + Kₐ = falling wood). Fourth, callers added meaningless calls to an already meaningful sequence and, in doing so, changed its meaning (e.g., B + Kₐ + H⁻ = neighbors). Fifth, callers added meaningful calls to an already meaningful sequence and, in doing so, refined its meaning (e.g., K + Kₐ = leopard; W⁺ + Kₐ = crowned eagle). We also found regularities in terms of call order. Boom calls, indicative of a nonpredation context, always preceded any other call types. H and K calls, indicators of crowned eagles or leopards, were always produced early in the sequence and were relatively more numerous if the level of threat was high.

In forest guenons, the single adult male of the group mainly vocalizes in response to predators and other significant external disturbances (5, 48, 53). Females are vocally active in social situations (44–46) and to predators to which they produce a diverse alarm call repertoire that encodes information about predator type and level of threat (47). Whether Campbell’s monkeys produced these calls to intentionally inform others about the event they have experienced cannot be decided with our data. Some observations suggest not. For example, it is puzzling that males produce loud calls in response to the thundering sound of falling trees, as all other group members will have perceived the event as well. The collapse of a large tree provides a significant danger to arboreal animals and fatalities are not unusual. Thus, calling males may have the urge to advertise their uninjured state and provide an acoustic focal point for scattered or disoriented group members, rather than attempting to inform others about the danger.
The core finding of this study is that the different calls produced by the males were usually given as part of structurally unique sequences and in context-specific ways. Context could be described in terms of event type, degree of threat, spatial relations within the group, and group movements. Some of these functions have already been shown for other guenons, particularly Diana monkeys (54) and putty-nosed monkeys (38, 39), but no previous study has revealed such a rich portfolio of context-specific calling behavior. In earlier playback experiments, series of H and K calls were meaningful to Diana monkeys, as these calls elicited their own corresponding eagle or leopard alarm calls (49). If artificially combined with pairs of B, however, Diana monkeys no longer responded to the H or K sequences, suggesting that this was due to the boom’s indication of a nonpredatory event. Although we have not found any natural B–H or B–K relations within the group, so there was no risk of confusion in terms of caller identification. If the male produced a loud call, the observer determined its likely cause among the following: (i) predator, presence of a leopard (Panthera pardus) or crowned eagle (Stephanoaetus coronatus); (ii) intergroup, presence of neighboring group (usually inferred by loud calls of their adult male); (iii) tree branch, crashing sound of falling tree or large branch; (iv) monkey alarm, Diana monkey eagle or leopard alarm calls (54); and (v) cohesion and travel, spatially separated far from the group (<30–50 m) or beginning of group travel after spatial separation.

The four habituated males were directly observed in dense forest vegetation for a total focal duration of 49 h (male 1, 11 months; male 2, 6 months; male 3, 15 months; male 4, 2 months). Scan animal samples (n = 4,425) were collected every 30 min, during which we recorded (i) the presence of other monkey species and (ii) the location within the group’s home range. We also collected ad libitum samples from all habituated and semihabituated groups during ~2,000 h total contact time. This enabled us notably to report on the response of males to their main predators, leopards (n = 3) and crowned eagles (n = 11).

Predator Experiments. As encounters with real predators were rare, we conducted a series of field experiments during which we presented predator models to the different males, either by positioning a visual replica of the predator or by broadcasting recordings of typical predator vocalizations (leopard growls, eagle shrieks). A custom-made eagle model and a custom-made leopard model were used for all experiments. Both manipulations have been used successfully before in other studies to simulate predator presence (e.g., 47, 54). We tested each of the four stimuli 10 times; that is, 40 trials total, on seven different males. The four semihabituated males were tested only once per stimulus category. Some of the habituated males were tested more than once, with subsequent trials spaced by at least 2 months (N leopard model; N eagle model; N leopard playback: male 1 = 3/2/3/2; male 2 = 1/2/1/1; male 3 = 2/2/2/3; male 4 was never tested as he arrived toward the end of the study).

Before each experiment, the following conditions had to be met: (i) the study group was aware of the presence of human observers for at least 30 min; (ii) no alarm calls were produced for at least 30 min; (iii) the predator model (or playback speaker) was positioned ahead of the group’s estimated travel direction ensuring that no associated monkey species could detect it first. One observer (K.O.) and one field assistant were necessary for these experiments. The observer walked with the group and recorded the male’s behavior, while the assistant operated the predator model. For eagle trials, the model or loudspeaker was positioned at an elevation of 2–3 m off the ground. For leopard trials, the model or loudspeaker was positioned on the ground. Eagle shrieks were recorded in the study area; leopard growls were purchased from the National Sound Archives, London. All acoustic stimuli were broadcast with a Sony WMDSC Professional Walkman connected to a NAGRA DSM speaker-amplifier with the amplitude level adjusted so that the calls sounded natural and could be clearly heard by the ears.

Within the predator context, we discriminated between degrees of urgency (i.e., likelihood of an attack) and compared the composition of sequences. Low-urgency situations were when the presence of a predator was revealed by acoustic cues (playback of predator calls, alarm calls by Diana monkeys); high-urgency situations were when the predator was visible (predator or by broadcasting recordings of typical predator vocalizations). This distinction was based on the results of another experimental study with forest African primates, which showed that call intensity was higher when a predator was detected in the visual than the acoustic domain (57).
Call Recordings and Analyses. All vocal responses were recorded using a Sony TCD-D100 DAT Walkman, a Sennheiser ME88 microphone, and a Lavaliere microphone for additional observer comments. All calls were digitized at a sampling rate of 44.1 kHz with 16 bits accuracy. A total of 224 male calling events were collected. In a previous study, we were able to reliably distinguish six different loud call types according to their acoustic structure: Boom (B), hok (H), krak (K), krak-oo (K-K), hok-oo (H-O), and wak-oo (W-J) (Table 1). Krak-oo and krak and hok-oo and hok calls were identical, apart from the optional -oo suffix. Krak-oo calls could be emitted singly; all other calls were given as part sequences of two or more calls. We determined the composition of each call sequence and related them to the eliciting context. To this end, we (i) identified all call types in the sequence, (ii) calculated the total number of each type produced, and (iii) determined the positioning of each call type in the sequence. Statistical analyses were conducted using the Statistics software package, using mainly nonparametric procedures.

ACKNOWLEDGMENTS. We thank the Office Ivoirien des Parcs et Réserve for permission to conduct research in the Tai National Park, the field assistants of the Tai Monkey Project for all their invaluable help during data collection, and N. Elizer, G. Emile, C. Blois-Heulin, V. Biquand, H. Bouchet, J. Fischer, M. Hausberger, and the GIS (groupe d’intérêt scientifique) "Cerveau, Comportement, Société" for support and discussions. This work was supported by the European Commission (Sixth Framework Programme, What it means to be human), European Science Foundation Eurocore programmes (Origin of Man Language and Languages), the Wissenschaftskolleg zu Berlin, the French ministry of foreign affairs (Egide), and the Centre Suisse de Recherches Scientifiques in Abidjan.