Brain & Language 120 (2012) 303–309

Contents lists available at SciVerse ScienceDirect

Brain & Language

journal homepage: www.elsevier.com/locate/b&l

Call combinations in monkeys: Compositional or idiomatic expressions?

Kate Arnold [*], Klaus Zuberbühler

Centre for Social Learning and Cognitive Evolution, School of Psychology, University of St. Andrews, St. Andrews, Fife KY16 9JP, UK

ARTICLE INFO

Article history:
Accepted 2 October 2011
Available online 26 October 2011

Keywords:
Language evolution
Syntax
Meaning
Combinatorial signal
Primate communication
Cercopithecus

ABSTRACT

Syntax is widely considered the feature that most decisively sets human language apart from other natural communication systems. Animal vocalisations are generally considered to be holistic with few examples of utterances meaning something other than the sum of their parts. Previously, we have shown that male putty-nosed monkeys produce call series consisting of two call types in response to different events. They can also be combined into short sequences that convey a different message from those conveyed by either call type alone. Here, we investigate whether 'pyow–hack' sequences are compositional in that the individual calls contribute to their overall meaning. However, the monkeys behaved as if they perceived the sequence as an idiomatic expression rather than decoding the sequence. Nonetheless, while this communication system lacks the generative power of syntax it enables callers to increase the number of messages that can be conveyed by a small and innate call repertoire.

1. Introduction

One of the greatest challenges for theories of human evolution is to provide a coherent account of how language has emerged from more ancestral forms of communication and cognition, and how this relates to animal communication. A hallmark of human language is its sheer expressive power, which is particularly striking when contrasted to all known primate vocal communication systems that are characterised by species-specific and highly restricted repertoires of calls. A widely held assumption is that syntax has emerged 'de novo', without any relevant precursors before humans diverged from the rest of the primates some six million years ago (e.g., Tallerman, 2005). Many animals produce monomorphic series of calls, and in some cases there is evidence that variations in series length are semantically meaningful features for listeners (e.g., Schel, Candiotti, & Zuberbühler, 2010; Stephan & Zuberbühler, 2008). Especially interesting examples are species that produce series composed of multiple call types in higher order structures. The most complex of these include the songs of some bird species (e.g., Catchpole & Slater, 1995; Gentner, Fenn, Margoliash, & Nusbaum, 2006; Marler & Slabbekoorn, 2004) and humpback whales (Payne & McVay, 1971; Suzuki, Buck, & Tyack, 2006). However, Hurford (2009) has argued that, unlike human language, the number of sequences is never larger than the number of elements that make them up. Equally important, there is no evidence that these combinations are semantically compositional. Rather, like simple, single-call utterances, they convey information about whole situations. Regarding their biological function, these utterances are sexually selected signals with the sole purpose of advertising territory ownership or the caller's competitive ability and suitability as a mating partner (Catchpole & Slater, 1995; Marler & Slabbekoorn, 2004; Tyack, 1981).

Within primates, the duet songs of some pair-bonded species, such as gibbons or titi monkeys, show significant complexity in terms of the number of call types they contain, with some indications that their organisation might be rule-based (Cäsar, Byrne, Young, & Zuberbühler, submitted for publication; Mitani & Marler, 1989; Robinson, 1984). Similar to bird song, this type of vocal behaviour appears to carry little meaning apart from advertising the presence of a bonded pair of residents to potential intruders and territorial neighbours. An interesting exception is the lar gibbons (*Hylobates lar*) that use their songs not only to advertise territory ownership but also in predator defence. Predator songs are made from the same basic song units as the more common duet songs, although the units are assembled in different ways (Clarke, Reichard, & Zuberbühler, 2006).

Whether the production of multiple calls is a strategy to generate redundancy (e.g., to increase the likelihood of successful perception) or to overcome the constraints of small signal repertoires is unknown for most species. For example, chimpanzees (*Pan troglodytes*) produce combinations of calls, and it has been suggested that some of them are linked to specific contexts (Crockford & Boesch, 2005), but more systematic work is required to test this hypothesis. There are some cases in which calls appear to act as contextual modifiers of others. Tamarins (*Sanguinus oedipus*) and capuchin monkeys (*Cebus olivaceus*), for example, produce combined utterances of alert and alarm calls, but as far as semantics are concerned, combining these calls seems to simply average the meaning of constituent calls (Robinson, 1984; Cleveland &

---

[*] Corresponding author. Fax: +44 1334 463042.
E-mail address: ka11@st-andrews.ac.uk (K. Arnold).

0093-934X/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.bandl.2011.10.001

Snowdon, 1982). In Campbell's monkeys (*Cercopithecus campbelli*), call sequences introduced by low-frequency 'boom' calls have been associated with a range of disturbances that do not involve the direct presence of a predator (Ouattara, Lemasson, & Zuberbühler, 2009a). In one playback experiment, boom calls were artificially added to two different alarm call series given to leopards and crowned eagles, respectively. This manipulation had significant effects on how listeners interpreted the meaning of the predator-specific alarm calls (Zuberbühler, 2002). In another study, it was shown that Campbell's monkeys produced additional acoustic variation by adding an invariant, suffixed, vocal unit to the basic alarm call types, which transformed them from predator-specific alarm calls to more generalised alarm calls (Ouattara, Lemasson, & Zuberbühler, 2009b). In sum, like animal songs, stereotypical call combinations occur in a number of primate species and it has been argued that this behaviour shows some parallels with phonemic combinations in human speech (Yip, 2006).

While such syntax-like signal production appears to be relatively infrequent, even in non-human primates, this is not the case for the perception of syntactically organised signals. In field experiments with Chacma baboons (*Papio cynocephalus*) listeners made social inferences when overhearing call exchanges between familiar group members. During social interactions, baboons produce specific vocalisations that reflect their relative rank (grunts, fear barks). In playback experiments, listeners responded weakly to simulated social interactions that complied with the social dominance hierarchy of the group (e.g., a dominant's grunts followed by a subordinate's fear barks; Cheney, Seyfarth, & Silk, 1995). If the sequence was reversed, so that listeners heard a subordinate's grunts followed by a dominant's fear barks, they responded strongly, especially if it involved members of different matrilines (Bergman, Beehner, Cheney, & Seyfarth, 2003). This suggests that baboons interpret vocal exchanges of others by taking into account the callers' relative social positions, which enables them to respond to simulated rank reversals. Similar findings have recently been reported for chimpanzees. Here, if listeners heard a low-ranking individual giving aggressor screams followed by victim screams of a higher-ranking one, they showed significantly more interest than to the reverse order (Slocombe, Kaller, Call, & Zuberbühler, 2010). Overall, these data suggest that non-human primates, and perhaps many other social animals, extract meaning by attending to sequences of signals given by multiple individuals. Whether or not listeners recruit the same cognitive processes when interpreting call sequences given by the <u>*same*</u> individual is an open question. During call exchanges between individuals, the meaning of individual calls remains largely unchanged (e.g., a 'fear bark' always indicates a socially subordinate caller). In contrast, in Campbell's monkey call sequences, the semantic content of individual calls appears to be mainly determined by their presence in a particular sequence (e.g., Ouattara et al., 2009a).

In this study, we address this issue by focussing on the semantic content of call combinations produced by putty-nosed monkeys. Like other guenon species, adult males' have a small repertoire of loud-calls, which carry over large distances and function in both intra- and intergroup communication. Two acoustically distinct loud call types 'pyows' and 'hacks' (Arnold & Zuberbühler, 2006a), are generally given as part of longer call series: 'pyow' series, 'hack' series, and 'transitional' series (a series of 'hacks' followed by a series of 'pyows'). These three series are given to a wide range of external events and there is evidence that listeners require contextual information to infer of the cause of the calls (Arnold & Zuberbühler, submitted for publication). However, 'pyows' and 'hacks' are also concatenated into short 'pyow–hack' (P–H) sequences that elicit a very different response from those elicited by any of the call series, i.e. they induce travel of the entire group (Fig. 1a; Arnold & Zuberbühler, 2006b, 2008). P–H sequences can be produced alone or inserted within any of the call series described above. In the latter case, they are distinguishable by significant pauses or by a transition back to 'pyows' (Fig. 1b).

While 'pyow', 'hack' or 'transitional' series resulted in increased vigilance or anti-predator behaviour, P–H sequences reliably elicited group travel without the need for additional contextual information. In a field experiment, we showed that call order, not variations in acoustic structure of the constituent calls, governed listeners' responses (Arnold & Zuberbühler, 2008). In putty-nosed monkeys, therefore, listeners attribute different meanings to the call strings, according to the different combinations of two call types.

The P–H call sequence provides a unique opportunity to explore whether animal call combinations approximate linguistic syntax beyond the phonological level (sensu Yip, 2006) for the following reason. Although P–H sequences are always short, the composition of the sequence varies significantly in terms of the number (between two and seven calls) and proportion (e.g., P P P P H vs. P H H H H) of its constituent calls.

In order to determine whether these compositional differences also differed in their semantic content, we carried out a field experiment in which we presented P–H sequences of different lengths and composition to females in a habituated group. If the P–H sequence is compositional, and the information conveyed by the overall sequence depends upon the relative contribution of each call type, then listeners should attend to the number of 'hacks' and 'pyows' in each sequence and respond differently. In previous work, we have show that 'pyow' series predominately cause listeners to orient their attention towards the caller, sometimes accompanied with some approach, while 'hacks' predominantly increase vigilance and inhibit movement (Arnold & Zuberbühler, 2008, submitted for publication). We thus predicted that P–H sequences composed of a large proportion of 'hacks' would result in significantly delayed travel over shorter distances compared to sequences containing a higher proportion of 'pyows'.

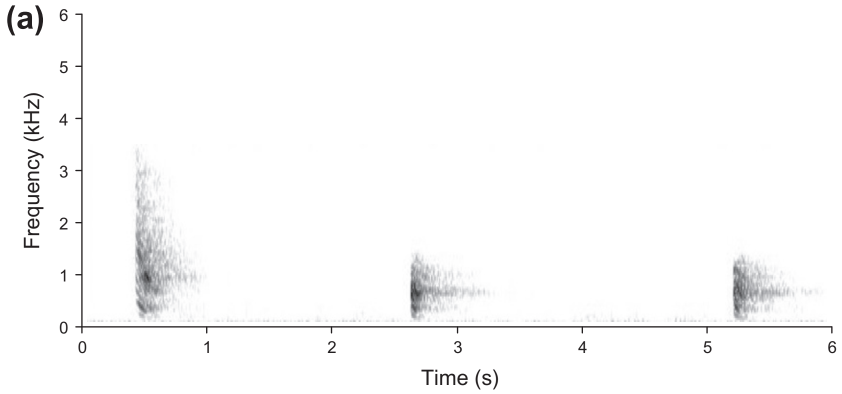
2. Methods

2.1. Study site and subjects

Field experiments were conducted in Gashaka Gumti National Park, Nigeria, between November 2008 and May 2009 by KA together with two field assistants. The study area consisted of a mosaic of primary semi-deciduous lowland rainforest and grassland near the village of Gashaka (7°20′N, 11°30′E). Putty-nosed monkeys live in one-male groups of up to 20 individuals, with 6–9 adult females and their offspring. Group density in the area has been estimated at 3–4 groups per km^2 (Dunn, 1993). One group of putty-nosed monkeys, which comprised one adult male, seven females and eight immature individuals during the period of study, had been followed on a daily basis since June 2007 and was habituated to human presence.

2.2. Adult male calls

Male putty nosed monkeys regularly produce two different kinds of loud calls, 'hacks' and 'pyows' (Struhsaker, 1970; Gautier-Hion, Colyn, & Gautier, 1999; Eckardt & Zuberbühler, 2004; Arnold & Zuberbühler, 2006a). Both vocalisations are loud and conspicuous, discrete call types that carry over one kilometre and can be distinguished by ear. Statistical analyses differentiating the acoustic structure of these call types have been presented elsewhere (Arnold & Zuberbühler, 2006a). 'Pyows' are individually distinctive and their acoustic structure differs subtly from male to male; 'hacks' on the other hand are not (Price, Arnold, Zuberbühler, & Semple, 2009).



(b)

P P H H H --* P P P P P P P P P P

H H H H H -- **P P H H H** -- H H H

H H H H H -- **P P H H H** --* P P P

Fig. 1. (a) Spectrographic illustration of a ‘pyow–hack’ sequence consisting of one ‘pyow’ and two ‘hacks’ given by a male putty-nosed monkey. (b) Typical patterns of insertion of ‘pyow–hack’ sequences within longer call strings – signifies a pause (* pause is optional where the sequence is followed by ‘pyows’).

2.3. Experimental protocol

From an extensive library of recordings of calls given by the group male we selected natural ‘pyow–hack’ sequences of different lengths and composition and edited them where necessary to control for overall length, using the software Cool Edit Pro 2.1 (Syntrillium Software), to produce the experimental stimuli. Stimuli consisting of six calls were of approximately the same duration. The stimuli consisting of two calls were necessarily shorter. We used only recordings of the group male’s calls as ‘pyows’ are individually distinctive (Price et al., 2009) and the broadcasting of the calls of another male at close proximity to the group would have been likely to affect subjects’ responses. Before each trial, the location of the male was established and his behaviour monitored by a field assistant throughout the trial. As soon as the male moved to the periphery of the group the experimenter selected one adult female, who was at least 50 m from the adult male, thus ensuring that she did not have the opportunity to observe his behaviour. A secondary requirement was that the focal subject was out of visual contact with other females to ensure that any subsequent behaviour was not guided by nearby individuals. Using walkie-talkies, a second field assistant was directed to a suitable location, where he set up the playback equipment consisting of a NAGRA DSM playback speaker connected to a portable CD player (Fig. 2).

Playback from this position ensured that the direction from which the female heard the stimulus was consistent with the known location of the male, and that other individuals within the group were not able to observe the equipment setup and learn the association between the equipment and the stimulus. Once the equipment was in place, one of the four stimulus types was broadcast and the locomotor behaviour of the target female was monitored for the next 20 min. Her initial position was read off a GPS unit at the time of the stimulus broadcast and at the end of the 20 min monitoring period to the nearest metre. The time that the subject began to move was also recorded.

The experiments were based on a within-subject design. On different days, one of the six known female subjects heard one of four playback sequences consisting of P–H sequences of different lengths and composition: (a) P–H–H–H–H–H, (b) P–P–P–H–H–H, or (c) P–P–P–P–P–H, (d) P–H. We conducted these 24 trials (6 × 4) in a randomised order. Trials were carried out at different times of the day and not more than once every 3 days. Trials were abandoned if the equipment setup was observed or if the group male responded with his own calls. Our dependent variables were the distance travelled during the next 20 min and the latency to travel following each playback stimulus.

2.4. Naturalistic observations

Observations of whole group responses to naturally occurring P–H sequences were also included in order to ensure that the results of the playback study were representative of natural behaviour. These data only include a measure of the distance travelled. Latency to travel was too difficult to determine accurately under natural conditions. For this purpose, a single habituated group of putty-nosed monkeys was followed for a total of 30 days between January and June 2005. Once located, the position of the estimated centre of mass of the group was recorded using a GPS unit. The group location was then recorded at 15-min intervals thereafter in order to determine the distance travelled. If the male produced loud calls, a field assistant recorded the group’s location, while KA recorded the sequence of calls, together with time that calling began. From that point on, the estimated centre of mass of the group was again recorded at 15-min intervals. Distance travelled was calculated for each uninterrupted 45-min block.

2.5. Statistical analysis

We conducted Friedman tests to compare the distance travelled and latency to travel following playbacks of differently composed P–H sequences. Where the results of Friedman tests were significant, we conducted exact Wilcoxon signed ranks post hoc tests in order to identify where significant differences lay. The relationship between the distance travelled and the composition of naturally occurring P–H sequences was investigated using Spearman’s rank correlations. All tests were two-tailed with α set at 0.05 except for post hoc comparisons. In these cases a Bonferroni correction was applied resulting in $\alpha = 0.017$ (0.05/3). Our second analysis strategy was to test the data with a generalised linear mixed model (GLMM). All analyses were carried out on SPSS 17.

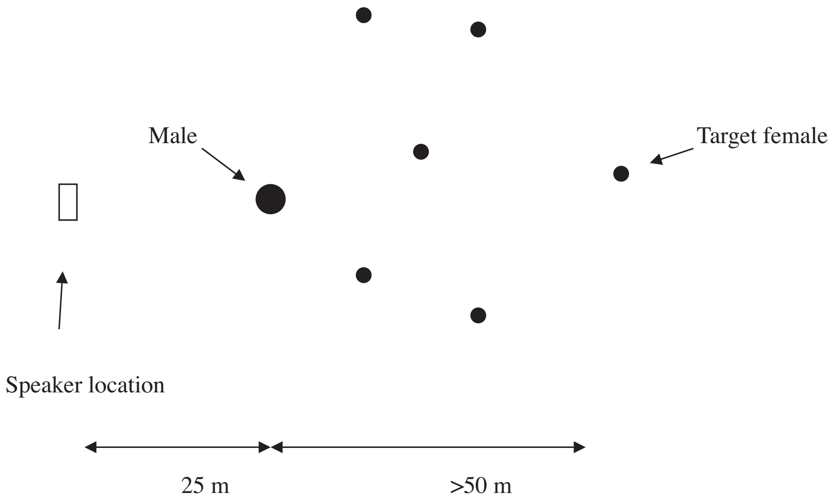


Fig. 2. Experimental setup with speaker location in relation to male, target female and other group members. The observer stayed with the target female and measured her locomotor response to the playback stimulus.

3. Results

We first analysed the distance travelled by entire groups, following naturally occurring P–H sequences. We recorded group movement after $N = 29$ instances of the male producing P–H sequences of various compositions. On $N = 3$ occasions the male produced two separate P–H sequences, one after the other. These observations were discarded from the analyses. The remaining $N = 26$ observations were P–H sequences of between two and seven calls and of variable composition.

Neither the absolute number of ‘hacks’, nor the proportion of ‘hacks’ in the P–H sequence resulted in systematic variations in the distance travelled (Spearman rank order correlation coefficient tests: number of ‘hacks’, $N = 26$, $r_s = 0.160$, $p = 0.435$; proportion of ‘hacks’, $N = 26$, $r_s = -0.123$, $p = 0.548$). However, there were indications of a positive relationship between the number of ‘pyows’ in a P–H sequence and the distance travelled by the group ($N = 26$, $r_s = 0.359$, $p = 0.072$; Fig. 3a) as well as the total number of calls of the sequence and the distance travelled ($N = 26$, $r_s = 0.318$, $p = 0.114$; Fig. 3b).

Males usually produced just one P–H sequence per occasion although sometimes they gave two ($N = 3$), one after the other, as indicated earlier. We found that the group moved significantly further following two P–H sequences than one (Mann–Whitney U test: $N_1 = 26$, $N_2 = 3$, $U = 11.5$, $p = 0.049$).

Although these observational results indicate that the number of ‘pyows’ per sequence is related to increased travel distance, this finding may be confounded by the sheer number of calls per sequence. In subsequent playback experiments, we addressed this potential confound by presenting short (2-call) or long (6-call) sequences with varying composition. Listeners usually responded to playbacks of different P–H sequences by moving towards the source of the calls. We found that the distance travelled differed according to the sequence type (median distance travelled following playbacks of PH = 9.5 m, PPPHHH = 28.5 m, PHHHHH = 22.5 m, PPPPPH = 24.0 m; $N_{PH} = N_{PPPHHH} = N_{PHHHHH} = N_{PPPPPH} = 6$; df = 1, $\chi^2 = 8.4$, $p < 0.001$; Friedman analysis of variance; Fig. 4a). However, post hoc tests revealed that this difference was due to listeners’ travelling shorter distances after hearing the short compared to the long P–H sequences (PH vs. PPPHHH, $z = -2.201$, $p = 0.031$; PH vs. PHHHHH, $z = -2.201$, $p = 0.031$; PH vs. PPPPPH, $z = -1.992$, $p = 0.063$; Wilcoxon signed ranks tests). Additional post hoc comparisons within the long P–H sequences indicated that differences

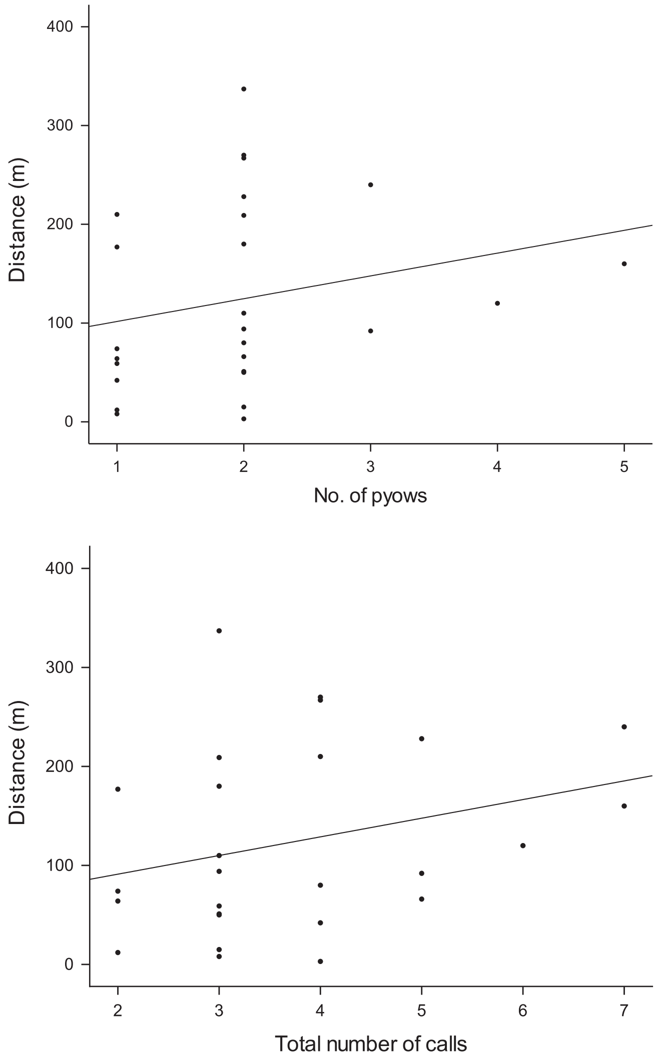
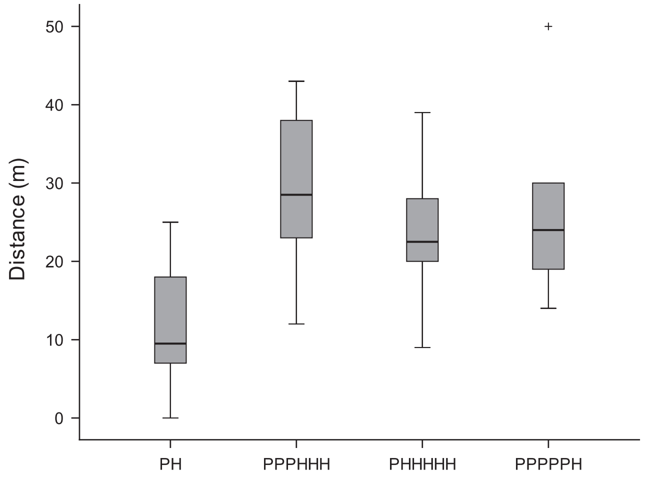


Fig. 3. The relationship between (a) the number of pyows, and (b) the total number of calls in naturally occurring P–H sequences and the distance travelled by the group.

(a) Distance travelled



(b) Latency to travel

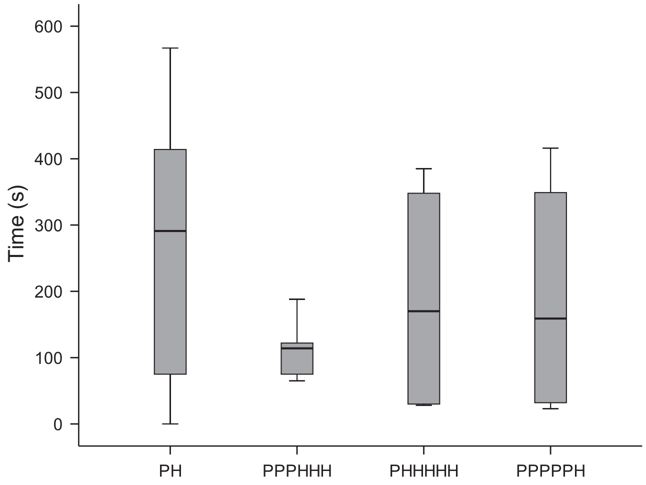


Fig. 4. The effect of different compositions of the P–H sequence on (a) the distance travelled, and (b) the latency to travel during the 20 mins following playbacks.

in composition, expressed in terms of the proportion of 'pyows' and 'hacks' within the sequence, had no effect on the distance travelled (PPPHHH vs. PHHHHH, $z = -1.153$, $p = 0.313$; PPPHHH vs. PPPPPH, $z = -0.526$, $p = 0.656$; PPPPPH vs. PHHHHH, $z = -0.105$, $p = 1$; Wilcoxon signed rank tests).

Our second analysis strategy was based on testing the data with a GLMM to determine whether distance travelled might be affected by an interaction between sequence length and the proportions of each call type within the sequence. However, this was not possible since there was no variance in the proportion of 'pyows' and 'hacks' in 2-call sequences. The main effects, however, were very much in line with the findings reported above (sequence length: $F(1,21) = 8.942$; $p < 0.01$; proportion of pyows: $F(1,21) = 0.301$; $p > 0.58$).

In terms of the latency to move, we also found differences according to the composition of the sequence (median latency following playbacks of PH = 291s, PPPHHH = 114s, PHHHHH = 170s, PPPPPH = 159s. Friedman analysis of variance: $N_{PH} = N_{PPPHHH} = N_{PHHHHH} = N_{PPPPPH} = 6$, df = 1, $\chi^2 = 20.16$, $p < 0.001$; Fig. 4b). However, none of the post hoc tests indicated a significant difference between P–H sequence types (Wilcoxon signed ranks test: PH vs. PPPHH, $z = -1.472$, $p = 0.188$; PH vs. PHHHHH, $z = -0.524$, $p = 0.688$; PH vs. PPPPPH, $z = -1.572$, $p = 0.156$; PPPHHH vs. PHHHHH, $z = -0.943$, $p = 0.438$; PPPHHH vs. PPPPPH, $z = -0.943$, $p = 0.438$; PPPPPH vs. PHHHHH, $z = -0.105$, $p = 1$).

4. Discussion

Our goal was to explore the possibility that an animal call combination may approximate linguistic syntax beyond the phonological level. One major objection is that it is difficult to assign meaning in any relevant conceptual sense to the male putty-nosed monkeys' calls since their use is rather broad, especially in the case of 'pyows'. Another is that the simple rule observed in putty-nosed monkey calls is a one-off and is not applied to other call types or other combinations. For example, there appears to be little difference, in either usage or response, between a series of 'hacks' and a 'transitional' series that begins with 'hacks' and 'ends' with 'pyows' (Arnold, Pohlner, & Zuberbühler, 2010). Nonetheless, it is possible that the two different call types that make up the sequence contribute to its overall meaning. In previous research we have shown that both 'pyows' and 'hacks' increase vigilance and orienting properties towards the caller. The two call types differ most strongly in the observed locomotor responses of listeners: while 'hacks' tend to inhibit movement (probably because they indicate eagle presence), 'pyows' tend to trigger movement towards the caller (probably to obtain additional cues about his gaze direction) (Arnold & Zuberbühler, submitted for publication). If monkeys attended to the compositional features of P–H sequences, then the degree of movement should have been correlated with the number of 'pyows' and 'hacks' in the sequence in either absolute or relative terms. Alternatively, listeners may simply learn that this particular, and fairly stereotypical, combination is associated with movement and need not unpack the message into its constituent elements. Our experimental results are consistent with the second hypothesis (Fig. 3).

While it is difficult to define precisely what 'pyows' and 'hacks' might mean, they do appear to have distinct functions when given alone or in series; series of 'hacks' are alarm calls associated with situations involving high urgency, particularly eagle attacks, while series of 'pyows' appear to function as attention-getters and are given in a wide variety of contexts (Arnold, Pohlner, & Zuberbühler, 2008; Arnold & Zuberbühler, 2006a, submitted for publication). If taken out of their native series and produced as part of a P–H sequence, do these two call types continue to contribute their native meaning or do they enter the sequence as semantically blank slates? We found two major effects that were relevant to this question.

Firstly and most importantly, the composition of the sequence (i.e. the proportion of 'pyows' and 'hacks' it contained) had no significant effect on listeners' responses for sequences with the same number of calls (Fig. 3). Under natural conditions, there was a tendency for the whole group to travel further after the male gave a P–H sequence, which contained a larger number of pyows, although this may have been confounded by the number of calls per sequence (Fig. 4).

Second, longer sequences resulted in larger distances being covered by listeners than shorter sequences. This effect was significant when comparing responses to experimental sequences consisting of two vs. six calls, regardless of composition (Fig. 3). Observations of whole group responses to sequences given under natural conditions showed a trend in the same direction (Fig. 4b), although the small sample size prevented meaningful conclusions. A relevant finding is that the few examples of the male repeating the sequence did result in significantly stronger responses. Overall, and like many animal communication systems, it appears that longer signals result in stronger responses, perhaps because they are easier to detect.

It may be noted that the distances travelled by females in response to playbacks were much shorter than those measured after naturally occurring P–H sequences. However, this is not surprising since, in natural situations, the male often leads a group progression after having instigated it with his vocal signals and females that are out of visual range of the male move in the direction of the call and eventually see the direction of travel taken by the rest of the group and follow them. However, in playback situations, the male had not called and remained in the same location. Playbacks, unsurprisingly, did not result in whole group travel. Focal females, however, were always out of visual range of the male and, as far as we could tell, other females. On hearing the male's calls they would begin to move in his apparent direction but generally stopped once they saw that other members of the group were not moving, hence the shorter observed travel distances.

A common strategy for signal identification and localisation is sensitivity to two or more spectral or temporal components (e.g., Capranica, 1965; Knudsen & Konishi, 1979). For example, tamarin monkeys respond to long calls containing two call types more readily than either call type alone. This has led to the hypothesis that nonlinear responses of neurons to combined signals, compared with linear responses to the individual elements, may underlie this behavioural selectivity (Ghazanfar, Flombaum, Miller, & Hauser, 2001). A further predication was that the central auditory system of species with long multi-syllabic vocal utterances contains more combination-sensitive neurons than other species, especially if the units of perception are combinations of calls or syllables. In the present case, longer P–H sequences may simply be easier to identify than shorter ones and are consequently responded to more readily.

There are some parallels between this communication system and the 'waggle dance' of honeybees, often regarded as one of the most complex examples of signal combinations. Foragers returning to the hive perform a dance, the angle of which, relative to the upward direction on the comb, indicates the direction of the food source relative to the sun. At the same time, the duration of each dance run is correlated with the distance to the food source (von Frisch, 1967; review by Dyer, 2002). Like the P–H sequence the 'intensity' of the signal (i.e. its duration) provides information about distance. However, there are also important differences. Different components of the bee dance encode different types of information (distance and direction) about past foraging events while the putty-nosed monkeys' P–H sequences refer to future behaviour. They signal the intention to move, and perhaps encode something about the intended distance, although in a more holistic way.

4.1. Linguistic analogies

To conclude, our results suggest that the P–H sequence is not compositional in any linguistic sense and if any parallel with human language can be drawn at all, it is with idiomatic phrases such as "kick the bucket", in which the meaning of the expression is not derived from the meaning of its constituent words but must be learned as a convention. 'Pyows' and 'hacks' in the context of a P–H sequence resemble free morphemes that have become affixed to each other and merged into a semantic unit (Saeed, 2003).

If this description is correct then this unique example of a meaningful animal call combination represents something of a syntactic dead end. There is no reason, in principle, why other distinctive combinations of pyows and hacks could not be used to convey other messages. So why is such potential not seized upon by natural selection (e.g., Novak, Plotkin, & Jansen, 2000)? It is unlikely that constraints are imposed by the ability of listeners to learn new sequences since the capacity to learn serial order and structural regularities within sequences has been demonstrated in non-human primates and birds (e.g., Endress, Carden, Versace, & Hauser, 2010; Gentner et al., 2006; Orlov, Amit, Yakovlev, Zohary, & Hochstein, 2006; Treichler, Raghanti, & Tilburg, 2003; Chen, Swartz, & Terrace, 1997). And the fact that listeners do not react to the 'hacks' within the series as they would when functioning as alarm calls indicates that they are able to dissociate calls from their usual function when given in the context of a particular sequence of calls. More complex communication, including the ability to innovate and link additional new sequences to specific contexts, may not be favoured by natural selection, assuming that the existing vocal repertoire of these monkeys is sufficient to deal with the main ecological and social challenges encountered by them. Our study also challenges the assumption that compound semantic units should arise as a solution to the problem of an increasingly unwieldy or poorly differentiated vocabulary of single phonemes (Jackendoff, 1999) since this male loud call repertoire of only two calls is the absolute minimum requirement for a combinatorial system to get off the ground. Hence, the constraints imposed by a small repertoire of calls may in fact have favoured the evolution of call combinations where a larger repertoire of signals offered an evolutionary advantage.

Acknowledgments

We thank the Nigerian National Park Services for granting permission to conduct research in Gashaka Gumti National Park and the Nigerian Conservation Foundation for help with logistics. We are grateful to Ali Tappare and Maikanti Hassan for their invaluable help in the field. We thank Robert Seyfarth for valuable comments and Roman Wittig for his statistical advice. This research has been funded by the Leverhulme Trust.

References

Arnold, K., Pohlner, Y., & Zuberbühler, K. (2008). A forest monkeys alarm calls to predator models. *Behavioral Ecology and Sociobiology, 62*, 549–559.

Arnold, K., Pohlner, Y., & Zuberbühler, K. (2010). Not words but meanings? Alarm calling behaviour in a forest guenon. In V. Sommer & C. Ross (Eds.), *Primates of Gashaka: Socioecology and conservation in Nigeria's biodiversity hotspot*. New York: Springer.

Arnold, K., & Zuberbühler, K. (2006a). The alarm calling system of adult male putty-nosed monkeys, *Cercopithecus nictitans* martini. *Animal Behaviour, 72*, 643–653.

Arnold, K., & Zuberbühler, K. (2006b). Semantic combinations in primate calls. *Nature, 441*, 303.

Arnold, K., & Zuberbühler, K. (2008). Meaningful call combinations in a nonhuman primate. *Current Biology, 18*, R202–R204.

Arnold, K. & Zuberbühler, K. (submitted for publication). Listeners require contextual information to disambiguate the cause of alarm calls in a forest primate. *Animal Cognition*.

Bergman, T. J., Beehner, J. C., Cheney, D. L., & Seyfarth, R. M. (2003). Hierarchical classification by rank and kinship in baboons. *Science, 302*(5648), 1234–1236.

Capranica, R. (1965). *The evoked vocal response of the bullfrog. Research monograph* (Vol. 33). Cambridge, MA: MIT Press.

Cäsar, C., Byrne, R., Young, R., & Zuberbühler, K. (submitted for publication). The alarm call system of wild black-fronted titi monkeys, *Callicebus nigrifrons*. *Behavioral Ecology and Sociobiology*.

Catchpole, C. K., & Slater, P. J. B. (1995). *Bird song: Themes and variations*. New York: Cambridge University Press.

Chen, S., Swartz, K., & Terrace, H. S. (1997). Knowledge of the ordinal position of list items in rhesus monkeys. *Psychological Science, 8*, 80–86.

Cheney, D. L., Seyfarth, R. M., & Silk, J. B. (1995). The responses of female baboons (*Papio cynocephalus ursinus*) to anomalous social interactions: Evidence for causal reasoning? *Journal of Comparative Psychology, 109*, 134–141.

Clarke, E., Reichard, U., & Zuberbühler, K. (2006). The syntax and meaning of wild gibbon songs. *PLoS One, 1*(1), e73.

Cleveland, J., & Snowdon, C. T. (1982). The complex vocal repertoire of the adult cotton-top tamarin (*Sanguinus oedipus oedipus*). *Zeitschrift fur Tierpsychologie – Journal of Comparative Ethology, 58*, 231–270.

Crockford, C., & Boesch, C. (2005). Call combinations in wild chimpanzees. *Behaviour, 142*, 397–421.

Dunn, A. (1993). The large mammals of Gashaka Gumpti National Park, Nigeria. Unpublished report to NCF, WWW-UK.

Dyer, F. C. (2002). The biology of the dance language. *Annual Review of Entomology, 47*, 917–949.

Eckardt, W., & Zuberbühler, K. (2004). Cooperation and competition in two forest monkeys. *Behavioral Ecology, 15*, 400–411.

Endress, A. D., Carden, S., Versace, E., & Hauser, M. D. (2010). The apes edge: Positional learning in chimpanzees and humans. *Animal Cognition, 13*, 483–495.

Gautier-Hion, A., Colyn, M., & Gautier, J. P. (1999). *Histoire naturelle des primates d'Afrique centrale*. Libreville, Gabon: Ecofac.

Gentner, T. Q., Fenn, K. M., Margoliash, D., & Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature, 440*, 1204–1207.

Ghazanfar, A. A., Flombaum, J. I., Miller, C. T., & Hauser, M. D. (2001). The units of perception in the antiphonal calling behavior of cotton-top tamarins (*Sanguinus oedipus*): Playback experiments with long calls. *Journal of Comparative Physiology A, 187*, 27–35.

Hurford, J. R. (2009). *The origins of grammar*. Oxford: Oxford University Press.

Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences, 3*, 272–279.

Knudsen, E. I., & Konishi, M. (1979). Mechanisms of sound localization in the Barn Owl (*Tyto alba*). *Journal of Comparative Physiology, 133*, 13–21.

Marler, P., & Slabbekoorn, H. (2004). *Nature's music: The science of birdsong*. New York: Academic Press.

Mitani, J. C., & Marler, P. A. (1989). Phonological analysis of male gibbon singing behaviour. *Behaviour, 106*, 20–45.

Novak, M. A., Plotkin, J. B., & Jansen, V. A. A. (2000). The evolution of syntactic communication. *Nature, 404*, 495–498.

Orlov, T., Amit, T. J., Yakovlev, V., Zohary, E., & Hochstein, S. (2006). Memory of ordinal number categories in macaque monkeys. *Journal of Cognitive Neurosciences, 18*, 399–417.

Ouattara, K., Lemasson, A., & Zuberbuhler, K. (2009a). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *PNAS, 106*, 22026–22031.

Ouattara, K., Lemasson, A., & Zuberbuhler, K. (2009b). Campbell's monkeys use affixation to alter call meaning. *PLoS One, 1*, e7808.

Payne, R., & McVay, S. (1971). Songs of humpback whales. *Science, 173*, 583–597.

Price, T., Arnold, K., Zuberbühler, K., & Semple, S. (2009). Pyow but not hack calls of the male putty-nosed monkey (*Cercopithecus nictitans*) convey information about caller identity. *Behaviour, 146*, 871–888.

Robinson, J. G. (1984). Syntactic structures in the vocalizations of wedge-capped capuchin monkeys, *Cebus olivaceus*. *Behaviour, 90*, 46–79.

Saeed, J. I. (2003). *Semantics* (2nd ed.). Oxford: Blackwell.

Schel, A. M., Candiotti, A., & Zuberbühler, K. (2010). Predator-deterring alarm call sequences in Guereza colobus monkeys are meaningful to conspecifics. *Animal Behaviour, 80*, 799–808.

Slocombe, K. E., Kaller, T., Call, J., & Zuberbuhler, K. (2010). Chimpanzees extract social information from agonistic screams. *PLoS One, 5*(7).

Stephan, C., & Zuberbühler, K. (2008). Predation increases acoustic complexity in primate alarm calls. *Biology Letters, 4*, 641–644.

Struhsaker, T. T. (1970). In J. R. Napier & P. H. Napier (Eds.), *Old world monkeys: Evolution, systematics and behaviour* (pp. 365–407). London: Academic Press.

Suzuki, R., Buck, J. R., & Tyack, P. L. (2006). Informational entropy of humpback whale songs. *Journal of the Acoustical Society of America, 119*, 1849–1866.

Tallerman, M. (2005). *Language origins: Perspectives on evolution*. Oxford: Oxford University Press.

Treichler, F. R., Raghanti, M. A., & Tilburg, D. N. V. (2003). Linking of serially ordered lists by macaque monkeys (*Macaca mulatta*): List position influences. *Journal of Experimental Psychology: Animal Behaviour Processes, 29*, 211–221.

Tyack, P. (1981). Interactions between singing Hawaiian humpback whales and conspecifics nearby. *Behavioral Ecology and Sociobiology, 8*, 105–116.

von Frisch, K. (1967). *The dance language and orientation of bees*. Cambridge, MA: Harvard University Press.

Yip, M. J. (2006). The search for phonology in other species. *Trends in Cognitive Sciences, 10*, 442–446.

Zuberbühler, K. (2002). A syntactic rule in forest monkey communication. *Animal Behaviour, 63*, 293–299.