Dwarf mongoose alarm calls:
investigating a complex non-human animal call

Katie Collier[a], Andrew N. Radford[b], Sabine Stoll[c,d], Stuart K. Watson[d,e], Marta B. Manser[a,d], Balthasar Bickel[d,e,*], Simon W. Townsend[a,d,e,f*]

[a]: Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
[b]: School of Biological Sciences, University of Bristol, 24 Tyndall Ave, Bristol BS8 1TQ, United Kingdom
[c]: Psycholinguistics Laboratory, University of Zurich, Plattenstrasse 54, 8032 Zurich, Switzerland
[d]: Center for the Interdisciplinary Study of Language Evolution (ISLE), University of Zurich, Switzerland
[e]:Department of Comparative Language Science, University of Zurich, Plattenstrasse 54, 8032 Zurich, Switzerland
[f]: Department of Psychology, University of Warwick, Coventry CV4 7AL, United Kingdom
*: Joint last authors

Abstract

Communication plays a vital role in the social lives of many species and varies greatly in complexity. One possible way to increase communicative complexity is by combining signals into longer sequences, which has been proposed as a mechanism allowing species with a limited repertoire to increase their communicative output. In mammals, most studies on combinatoriality have focused on vocal communication in non-human primates. Here, we investigated a potential combination of alarm calls in the dwarf mongoose (*Helogale parvula*), a non-primate mammal. Acoustic analyses and playback experiments with a wild population suggest: i) that dwarf mongooses produce a complex call type (T_3) which, at least at the surface level, seems to comprise units that are not functionally different to two meaningful alarm calls (aerial and terrestrial); and ii) that this T_3 call functions as a general alarm, produced in response to a wide range of threats. Using a novel approach, we further explored multiple interpretations of the T_3 call based on the information content of the apparent comprising calls and how they are combined. We also considered an alternative, non-combinatorial, interpretation that frames T_3 as the origin, rather than the product, of the individual alarm calls. This study complements previous knowledge of vocal combinatoriality in non-primate mammals and introduces an approach that could facilitate comparisons between different animal and human communication systems.

Background

Communication plays an essential role in the social lives of many species [1–4], with considerable interspecific variation both in the modality used (e.g. auditory, visual, olfactory) and in complexity [5]. Communicative complexity has long been defined in several ways: for example, by the number of structurally and functionally distinct elements or the amount of bits of information, with the presence of more elements or bits representing more complex systems [6,7]. Recently, one measure of communicative complexity in the vocal domain – the capacity to combine calls together into larger structures – has received increasing attention. Comparative and theoretical work suggests that combining calls not only serves to increase the communicative output of a species but, compared to the creation of new calls, does so in a more efficient way [8,9] and with a reduced error risk for the receiver [10,11].

To date, the majority of research on signal combinations in non-human vocal communication has focused on primate species, not least because of their close phylogenetic relationship to humans and thus the potential to shed light on the evolution of our own highly combinatorial communication system. Combinatorial capacities have, for example, been demonstrated in both primate alarm and long calls (black-fronted titi monkeys, *Callicebus nigrifrons* [12,13]; Bornean orang-utans, *Pongo pygmaeus wurmbii* [14]; putty-nosed monkeys, *Cercopithecus nictitans* [15,16]; white-handed gibbons, *Hylobates lar* [17]) and their social calls (chimpanzees, *Pan troglodytes* [18]; bonobos, *Pan paniscus* [19,20]; red-capped mangabeys, *Cercocebus torquatus* [21]; Diana monkeys, *Cercopithecus diana diana* [22]). One combinatorial signal in Campbell's monkeys (*Cercopithecus campbelli campbelli*) has received particular attention, primarily due to its structural similarity with compositionality in human language where meaning-bearing units (e.g. words) are combined together into larger meaningful structures [23,24]. Specifically, Campbell's monkeys have been shown to affix an acoustically distinct '-oo' unit to their predator-specific alarm calls [25,26]. The addition of this

64 affix changes the meaning of different alarm calls in a predictable way, from specific to general, and

65 has therefore been interpreted as a rudimentary combinatorial or even compositional system [27–29].

66   Despite an emphasis on primates, recent research suggests similar combinatorial capacities

67 are also present in taxa more distantly related to humans. For example, two bird species have been

68 demonstrated to produce remarkably similar combinations of calls. Both pied babblers (*Turdoides*

69 *bicolor*) and Japanese great tits (*Parus minor*) combine alert vocalisations (used to indicate threats)

70 with a recruitment call (used to recruit conspecifics in a variety of events) into a larger structure when

71 encountering threats, such as snakes, that require recruitment [30,31]. Playback experiments have

72 confirmed that these call combinations are meaningful to receivers, conveying information on both

73 the context and the required action [30,31]. There also exists intriguing, detailed observational data

74 documenting call combinations in non-primate mammals (banded mongooses, *Mungos mungo* [32];

75 meerkats, *Suricata suricatta* [33]). In comparison to birds and primates, however, experimental

76 verification of the structure and function of these combinations is still needed (though see [34] for an

77 example in dingos, *Canis familiaris dingo*). If we are to capture the complexity of animal vocal

78 communication systems, data on the production and perception of call combinations are required

79 across a wide range of species and taxa. Such data are particularly important for understanding the

80 role that combinatoriality might play in facilitating the emergence of complex communication

81 systems [35]. Here, we aim to further existing knowledge by experimentally investigating

82 combinatorial-like structures in the alarm-call system of a non-primate mammal, the dwarf mongoose

83 (*Helogale parvula*).

84   Dwarf mongooses are a highly social, small carnivore species from Eastern and Southern

85 Africa. They live in groups of up to 30 individuals [36], composed of a dominant pair and subordinate

86 individuals of both sexes who can be related or unrelated to the dominant male and female [37]. They

87 forage for insects and small vertebrates as part of a group. During foraging sessions, individuals often

88 perform sentinel behaviour, standing in an elevated position and alerting the rest of the group to

89 threats by producing alarm calls [38]. Among other calls, dwarf mongooses produce two predator-

90 specific alarm calls upon detecting aerial and terrestrial predators [39]. A third alarm call, type 3 (T_3),

91 appears to be a combination of aerial and terrestrial alarm calls [39] (Figure 1), with the structure

92 resembling an aerial alarm call (hereafter $T_{3.1}$) always preceding the structure resembling a terrestrial

93 alarm call (hereafter $T_{3.2}$). Previous observations indicate that T_3 functions as a general alarm call (i.e.

94 a call given to a wide range of disturbances which contains no specific information about the eliciting

95 context [40]), and can occur in ambiguous predation contexts [39], potentially suggesting that the

96 meaning of T_3 is related to the meaning of its parts. However, an experimental demonstration that the

97 T_3 alarm call functions to communicate general threats is lacking.

98   Building on prior work, here we used field-based playback experiments and acoustic analyses

99 to investigate further the function and acoustic structure of T_3 alarm calls and to determine the extent

100 to which the overall meaning of the combination is derived from its parts. First, if T_3 serves as a

101 general alarm call, we expected to observe general anti-predator behaviours in response to its

102 playback. Second, if T_3 has this function by virtue of being, at least at the surface level, a combination

103 of an aerial and a terrestrial call: i) the acoustic structure of T_3's component parts should be similar to

104 the acoustic structures of the corresponding alarm call types; and ii) playbacks of the different parts of

105 T_3 in isolation should reveal similar responses to the aerial and terrestrial alarm calls.

106

107 **Methods**

108 *Study site and population*

109 Our research was carried out between November 2014 and June 2015, and in January–February 2016,

110 as part of the long-term Dwarf Mongoose Research Project. Subjects were adult (>1 year) dwarf

111 mongooses living in their natural habitat located on Sorabi Rock Lodge, South Africa [41] and

112 belonged to seven wild but habituated groups composed of 6 to 15 individuals (mean group size=11).

113 Individuals were habituated to close observations, allowing sound recordings from 1–3 m and detailed

114 data-collection from field-based playback experiments [39,41]. All mongooses were individually

115 identifiable by small blonde dye-marks or distinguishable features such as scars [41].

116

117

118

119   *Alarm call collection and acoustic analysis*

120   Alarm calls for acoustic analysis were collected through employing the same methods as in previous

121   work [39]. Dwarf mongooses were followed during two daily sessions, one in the morning and

122   another in the evening, during which we recorded, *ad libitum,* all alarm calls produced. When

123   possible, the eliciting stimulus, the mongooses' response and the caller's identity were noted. To

124   collect additional alarm calls, in particular those produced in the presence of terrestrial predators

125   (none of which we obtained naturally), we carried out predator presentations. We used a domestic dog

126   (husky crossbreed, approx. 60 cm at the shoulders) on a leash as a substitute terrestrial predator and a

127   helium balloon to simulate an aerial predator [39].

128   We performed an acoustic analysis to compare the different types of alarm calls emitted and

129   to determine whether T_3 is structurally a combination of an aerial and a terrestrial alarm. We first

130   visualized the calls using Praat version 5.3.85 (www.praat.org) and selected good-quality alarm calls

131   with a high signal-to-noise ratio for the analysis. Using the acoustic program Luscinia [42], we then

132   extracted several temporal and spectral parameters from these calls (Table 1). We did this for each of

133   the natural alarm call types (aerial, terrestrial and T_3) but also for the first and second halves of T_3 ($T_{3.1}$

134   and $T_{3.2}$), respectively resembling an aerial and a terrestrial alarm call.

135

136   *Playbacks*

137   Following on from previous work investigating the contexts in which the different dwarf mongoose

138   alarm-call types are produced [39] (results summarised in Table 2), we aimed to clarify T_3's status as a

139   general alarm call by considering the responses to playbacks of three natural alarm calls: aerial,

140   terrestrial and T_3. For natural alarm calls, we performed a total of 18 playbacks each for aerial and

141   terrestrial calls (to 17 individuals belonging to seven groups) and 15 playbacks for T_3 alarms (to 14

142   individuals belonging to seven groups; in all cases, the individual that received two playbacks of the

143   same call type was tested in two different field seasons).

144   To test whether subjects perceived the T_3 alarm call as a combination of aerial and terrestrial

145   alarms, we carried out paired playbacks of three sets of stimuli: i) T_3 and artificial T_3 alarms (T_{3art}),

146　created by sequentially concatenating the recordings of individually produced aerial and terrestrial

147　calls; ii) aerial and $T_{3.1}$; and iii) terrestrial and $T_{3.2}$ alarms. The aim was to conduct the paired

148　playbacks to 10 focal mongooses, belonging to seven different groups, but set (i) could only be played

149　back to eight individuals. For all playbacks, we selected calls with a good signal-to-noise ratio as

150　stimuli, providing 15 exemplars of aerial, 12 of terrestrial, and nine of T_3 alarm calls. Each stimulus

151　consisted of a single call recorded from a foreign group to avoid the focal individual hearing its own

152　alarm call during the experiment. We played back the alarm calls from a height of about 1 m to

153　simulate a call from an individual acting as a sentinel (raised guard) [41]. We implemented playbacks

154　when the focal mongoose was foraging in the open and its response was filmed from 3 to 5 m using a

155　handheld camcorder (Canon Legria HF R506; Cannon Inc., Tokyo, Japan). All behavioural

156　measurements were taken from the resulting videos. We noted the focal individual's immediate

157　response to the playback and, in line with previous work [39], scored its strength according to the

158　activity or energy required by the different responses, which also corresponds to the scale of

159　disruption that the response causes to mongoose foraging activity. Specifically, we scored 1=no

160　reaction (no visible change in behaviour); 2=vigilance (mongoose paused foraging and scanned the

161　area horizontally); 3=moved (mongoose started moving but stopped short of cover); or 4=ran for

162　cover (mongoose moved quickly to the nearest bush or rocks). We also determined the focal

163　individual's latency to relax; that is, latency to resume foraging or engage in auto- or allo-grooming.

164　Furthermore, we noted whether, within the minute following playback, the mongoose engaged in

165　additional anti-predator behaviours: looking-up behaviour (i.e. looking at the sky), which may

166　facilitate the detection of aerial predators; or initiating a sentinel bout, which could allow subjects to

167　detect any type of predator. To ensure accurate coding of the videos, 15 randomly selected videos

168　(26% of trials) were blind-coded by a second naïve observer. Interobserver analyses suggest a reliable

169　agreement between observers: (% agreement varied between 73% and 86% for the different categories

170　of behaviours; strongest reaction (r=0.89, p<0.001), looking up (adjusted Kappa: 0.73 [CIs: 0.19,

171　0.97) and becoming a sentinel (adjusted Kappa: 0.6 [CIs: 0.04, 0.91] [43,44]. Experimental trials only

172　took place if no conspecific or heterospecific alarm calls had been heard within the last 10 min and

173　the mongooses were showing no signs of alarm or arousal from a previous event (predator encounter

174   or intergroup interaction). At least 1 h separated two successive playbacks, with a maximum of three

175   playbacks per session (morning or afternoon). All stimuli were presented in a random order.

176

177   *Statistical analysis*

178   a) Acoustic analysis

179   We compared the three natural call types, obtained from five different groups, to each other (N=7

180   calls per group per type), as well as aerial and terrestrial alarm calls to $T_{3.1}$ and $T_{3.2}$ respectively (aerial

181   and $T_{3.1}$: N=7 calls per group per type; terrestrial and $T_{3.2}$: N=10 calls per group per type), using the

182   measured acoustic parameters (Table 1). We started by removing any collinear parameters, as

183   determined by their variance inflation factors (VIF). We calculated VIFs for all parameters and

184   discarded the parameter with the highest VIF and then repeated these steps until all remaining

185   parameters had VIFs with values lower than 10 and therefore should not be collinear [45]. We then

186   used the remaining parameters to run a Discriminant Function Analysis (DFA). Given that multiple

187   calls obtained from the same group contributed to the dataset, we implemented permutated DFAs

188   (pDFA) using a function provided by R. Mundry. Unlike conventional DFAs, pDFAs allow for

189   repeated measures due to multiple recordings of an individual or group and do not return inflated p-

190   values [46]. Ideally, we would also have controlled for potential repeated measures at the individual

191   level, but this was not possible due to difficulties reliably identifying callers on a regular basis. All

192   analyses were performed using R version 3.2.1 [47] with the packages usdm [48] and MASS [49].

193

194   b) Playbacks

195   To investigate the strength of reaction when hearing the three different natural alarm calls, we used a

196   Cumulative Link Mixed Model (CLMM), fitting stimulus type as a fixed effect and individual nested

197   within group as a random effect. When a significant result was returned, we carried out post-hoc

198   pairwise CLMMs between the treatments (aerial vs terrestrial, aerial vs T_3, terrestrial vs T_3) and p-

199   values were adjusted for multiple testing using Bonferroni's correction. Models of the same format

200   were used to compare the strength of reaction between paired stimuli (aerial and $T_{3.1}$, terrestrial and

201    $T_{3.2}$, T_3 and T_{3art}).

202   To compare latencies to relax in response to the three different natural alarm calls, we carried

203   out a Linear Mixed Model (LMM) with stimulus type as a fixed effect and individual nested within

204   group as a random effect. Inspecting plots of the model residuals showed that our data did not violate

205   the assumptions of linearity, homoscedasticity and normality of the residuals. However, this was not

206   the case for the data from the paired playbacks, so we used non-parametric Wilcoxon signed-rank

207   tests to compare latencies to relax in this case.

208   To test whether the expression of the additional anti-predator behaviours (looking-up and

209   acting as a sentinel) differed in response to different playback stimuli, we performed Generalized

210   Linear Mixed Models (GLMMs) with a binomial family and a logit link function with stimulus type

211   as a fixed effect and individual nested within group as a random effect. This was done for all

212   playbacks except in the case of looking-up behaviour in paired-playback set (iii), as this behaviour

213   was not expressed in reaction to the stimuli used. For all the models, p-values were obtained using

214   likelihood ratio tests comparing full models, including all the explanatory variables, to reduced

215   models including the same explanatory variables but without the variable of interest. Analyses were

216   performed using R [47] with the packages ordinal [50] and lme4 [51].

217

218   **Results**

219   *Acoustic analysis*

220   The three natural alarm call types were distinguishable by the measured acoustic parameters (pDFA:

221   N_{calls}=105, p=0.002, percentage correctly cross-classified=82%). Aerial alarms and the first element of

222   T_3 ($T_{3.1}$) could not reliably be distinguished from each other by acoustic parameters alone (N_{calls}=70,

223   p=0.091, percentage correctly cross-classified=68%), whereas terrestrial alarms and the second

224   element of T_3 ($T_{3.2}$) could be discriminated (N_{calls}=100, p=0.026, percentage correctly cross-

225   classified=94%).

226

227   *Function of T_3 alarm calls*

228   The strength of reaction by dwarf mongooses to playbacks of natural alarm calls depended on alarm-

229   call type (CLMM: χ^2_2=6.88, p=0.03; Figure 2). Whilst we have previously shown that subjects reacted

230 differently to aerial and terrestrial alarms [39], Bonferroni-corrected post-hoc pairwise comparisons of

231 the current data indicated that reaction strength was not significantly different in response to aerial

232 and T_3 alarm calls (χ^2_1=1.27, p_{adj}=0.78) and to terrestrial and T_3 alarm calls (χ^2_1=2.01, p_{adj}=0.48). The

233 absence of differences in reaction strength to T_3 and aerial or terrestrial calls, in addition to its

234 previously defined use in multiple and ambiguous predator contexts [39], is highly suggestive of T_3's

235 status as a general alarm call. In addition, there was no significant difference in latency to relax

236 (LMM: χ^2_2=1.90, p=0.39) or sentinel behaviour (χ^2_2=0.28, p=0.87), in response to playback of the

237 different natural alarm calls. Furthermore, there was no significant difference in subsequent looking-

238 up behaviour (GLMM: χ^2_2=4.98, p=0.083).

239

240 *Playback of paired natural and experimentally modified alarm calls*

241 We found that the strength of response did not differ significantly to T_3 and T_{3art} (CLMM: χ^2_1=0.22,

242 p=0.26) or to aerial and $T_{3.1}$ (χ^2_1=3.06, p=0.08). Furthermore, despite differences in acoustic structure

243 between $T_{3.2}$ and terrestrial alarm calls, we also found no difference in strength of response between

244 these calls (χ^2_1=1.25, p=0.26; Figure 3). One possible explanation for $T_{3.2}$'s differing acoustic structure

245 could therefore be a co-articulation mechanism, in which the properties of a sound are modified by the

246 influences of adjacent sounds [52]. There was, in addition, no significant difference in latency to relax

247 (Wilcoxon, T_3 and T_{3art}: v=16.5, N=8, p=0.51; aerial and $T_{3.1}$: v=21, N=10, p=0.73; terrestrial and $T_{3.2}$:

248 v=12, N=10, p=0.83), in looking-up behaviour (GLMM, T_3 and T_{3art}: χ^2_1=0, p=1; aerial and $T_{3.1}$: χ^2_1=0,

249 p=1) or in sentinel behaviour (GLMM, T_3 and T_{3art}: χ^2_1=0.40, p=0.53; aerial and $T_{3.1}$: χ^2_1=1.41,

250 p=0.23; terrestrial and $T_{3.2}$: χ^2_1=0.43, p=0.51) between the pairs of playbacks.

251 Given the nature of p-values and the relatively small sample sizes, we wished to exercise

252 caution when inferring a lack of effect from non-significant results. We therefore reran our analyses

253 using Bayesian mixed-effects models, which are able to estimate parameters more accurately from

254 smaller sample sizes than maximum-likelihood approaches [53]. This analysis supported the

255 outcomes of our original frequentist analysis in that there was no strong evidence of an effect of

256 playback type on any behavioural response except terrestrial vs aerial alarm calls (See Electronic

257 Supplemental Material for details).

Discussion

Our observational and experimental results indicate that dwarf mongooses produce a complex call type (T_3) that, at least at the surface level, seems to comprise units that are not clearly functionally different from two meaningful alarm calls (aerial and terrestrial). Furthermore, the strength of reaction to playbacks suggests that the structure of the T_3 call appears meaningful to dwarf mongooses, with subjects reacting to it in an intermediary, or generalised, way: they respond with a mix of behaviours, compared to the two distinct alarm calls, rather than with a novel qualitatively different response as in [15]. We consider both combinatorial and non-combinatorial hypotheses for the origin of T_3.

One possible interpretation of the T_3 alarm call is that it represents a combination of two different alarm calls and that by recombining existing meaningful calls from the repertoire in transparent ways, dwarf mongoose are capable of communicating related, yet subtly different information [26,30,31]. Decomposing how the precise information content of two predator specific alarm calls gives rise to a more general alarm call is, however, non-trivial. Here we analyse the calls within a linguistic framework, adopting technical notions from language sciences. It is important to highlight that in doing so, we do not mean to imply the same cognitive mechanisms are at play as when combinations are processed in human language. Rather, it represents one method to assess the precise use conditions of animal calls. Specifically, we considered different possible interpretations of the T_3 alarm call depending on the potential contexts of use and the associated informational content of the distinct aerial and terrestrial alarm calls and the ways in which they could be combined. For example, in line with previous findings in two bird species [30,31], one potential interpretation would be that the alarm components denote their respective predator type and T_3 would act as a "conjunction" [54], denoting the presence of both predators simultaneously. This interpretation, however, is incongruent with the data because this conjunction of events was never observed, yet T_3 was commonly emitted (20% of all alarm calls recorded) [39]. Moreover, we would expect a stronger reaction to an alarm call denoting the presence of two types of predator instead of one, which was not the case.

Several other interpretations can be considered. In the first, the alarm components could indicate the behaviour that receivers should perform (run for cover; vigilance). In which case, T_3

286 would convey something akin to "run for cover and be vigilant". Whilst possible, this interpretation of

287 the individual calls is problematic, which in turn makes this interpretation of T_3 unlikely. That is

288 because the experimental data suggest dwarf mongooses sometimes become vigilant in response to an

289 aerial alarm or run for cover when hearing a terrestrial alarm (Figure 2), implying that these calls do

290 not denote the behaviour receivers should perform. Furthermore, an additional cognitive step would

291 be required on the caller's part as, once the caller has identified the type of threat (e.g. aerial or

292 terrestrial predator), it would then have to convert predator type into the appropriate reaction for

293 receivers to perform (e.g. run for cover, be vigilant) before emitting the alarm call.

294   In another possible interpretation, the aerial and terrestrial alarm calls could refer respectively

295 to aerial and terrestrial predators, and T_3 refers to an aerial *or* a terrestrial predator. This interpretation

296 would represent a "disjunction", in which two units are combined with at least one context being

297 appropriate or "true", but not necessarily both (e.g. *turn left **or** right*) [54]. This interpretation of T_3

298 would be very inefficient, indicating two very distinct forms that a threat could take, leaving receivers

299 uncertain as to the exact nature of the danger. In these circumstances, as a disjunction does not

300 provide specific information on the type of threat but only that danger is present, we would predict

301 that mongooses would alternate between vigilance (to detect a terrestrial threat) and looking up at the

302 sky (to detect an aerial threat) to establish first what the probable threat is before engaging in any

303 potentially dangerous anti-predator behaviour that might be detrimental to survival. Yet this is not

304 what we observed: mongooses only occasionally looked at the sky after hearing a T_3 alarm call (2/15

305 trials) and this was always after running for cover. Nevertheless, the disjunction analysis does seem to

306 capture something genuine about the function of this combination and its use as a general alarm call:

307 enumerating the alternatives (here: raptor, terrestrial predator) as a way of generalizing across them

308 (danger).

309   Following on from this, an additional potential interpretation is that T_3 calls would have a

310 similar structure to what are termed 'listing compounds' in human language. Listing compounds also

311 define a context by enumerating the possible alternatives. An example from English is "pass–fail", as

312 in a "pass–fail exam", which details all possible contexts or outcomes of this type of exam. The

313   critical difference between a listing compound ("a pass–fail exam") and an explicit disjunction ("an

314   exam in which you can pass or in which you can fail") in language is that a compound is interpreted

315   as a whole and points to a single specific context, in this case an established type of exam, while the

316   disjunction rests on interpreting each statement on its own ("you can pass it", "you can fail it")

317   followed by complex logical operation that combines the two by "or". Under this analogy with listing

318   compounds in language, T_3 calls would signal a more general danger context, defined by listing its

319   alternatives. This interpretation reflects the fact that the component calls indeed signal disjoint

320   contexts, but it is at the same time consistent with the fact that a combination refers to a single context

321   and is not ambiguous between two contexts.

322   Whilst general alarm calls are common in non-human animals [55], it is unclear why dwarf

323   mongooses use a combination of independently occurring calls, rather than a single call to fulfil this

324   general alarm function. Interestingly, a similar phenomenon also exists in Campbell monkeys who

325   have been demonstrated to use call combinations (e.g. krak-oo) in more general threat situations, as

326   opposed to the single calls (e.g. krak) used in response to a specific predator type [26]. In line with

327   theoretical modelling work, it is possible that in dwarf mongooses, each of the individual alarms

328   accomplishes a specific function, leading to the compounding of calls to communicate additional

329   information [35]. It is worth noting that whilst aerial and terrestrial calls are primarily associated with

330   a specific behavioural response (run for cover and vigilance respectively) this is more of a

331   probabilistic rather than a deterministic relationship. It could be, therefore, that the intermediate

332   response elicited by T_3 is, instead, a result of receivers probabilistically inferring the referent of two

333   separate alarm calls sequentially leading to a more noisy behavioural response.

334   Alternatively, rather than the T_3 call being a combination of two independently occurring calls, an

335   equally plausible interpretation is that the T_3 call represents a stand-alone, holistically meaningful call

336   from which the more specific aerial and terrestrial calls are derived. Such an analysis is particularly

337   attractive as it is potentially simpler (in an evolutionary sense): if T_3 is in fact a single, albeit

338   acoustically complex, call, then there is no need to explain why a combination of calls, rather than a

339   single call type, would be used to signal a general threat. This alternative scenario also has important

340   evolutionary implications. To date, the majority of work focusing on combinatoriality in animal

341 communication has posited that it serves to expand the vocal repertoire, particularly in species that are

342 constrained in their vocal production [15,30]. However, it may well be that repertoire size is

343 expanded, not through combinatoriality, but instead by decomposing complex calls into smaller, more

344 specific parts.

345   It is evident that more research is ultimately needed. Though reanalysis of our data within a

346 Bayesian framework broadly corroborates our findings, increasing the power with a larger sample size

347 would be important. Furthermore, additional playback experiments, in particular, could be a fruitful

348 way to begin to disentangle the competing hypotheses regarding T_3's origins and meaning. For

349 example, in line with the proposal that T_3 represents a third distinct call type (as opposed to a

350 combination of individual calls), a playback experiment reversing the order of units structurally

351 equivalent to aerial and terrestrial calls should lead to a loss in behavioural response, as the call is

352 never produced in this way. If, on the other hand, T_3 is a combination of independent alarm calls and

353 derives its meaning, in some way, from these individual parts, dwarf mongoose should still be able to

354 process the meaning of T_3, irrespective of the order of its components. Moreover, manipulations to the

355 duration between calls, or indeed simulating T_3 from callers in two different locations, could shed

356 additional light on whether T_3 is a bona fide, simple, syntactic structure, as opposed to an unrelated

357 sequence of two calls that happen to fall adjacent to each other.

358

359 *Conclusions*

360 Our study offers a new example of a complex call structure (T_3) in mammals that superficially

361 resembles a combination of two individual calls (aerial and terrestrial alarm calls). This research in

362 dwarf mongooses not only complements previous research on combinatoriality in animal

363 communication but also helps shed light on the phylogenetic distribution of this phenomena. This, in

364 turn, will ultimately help inform our knowledge about the social and environmental factors promoting

365 such vocal complexity and, more specifically, the exact type of combinatoriality employed by a

366 species (e.g. compounding, disjunction, conjunction). Further work is still necessary to rule out other

367 potential explanations – namely that T_3 is actually the source of the individual calls, rather than being

368   a product of combining two calls. Such an analysis represents an intriguing avenue for future research

369   in dwarf mongoose communication but also in animal combinatoriality in general. Lastly, to our

370   knowledge, this is the first attempt at analysing a non-human animal call combination based on the

371   possible meanings of the combination's individual components and the different ways in which they

372   can be combined. Such an approach is key to unpacking not only the similarities and differences

373   between combinations in various animal communication systems, but also between such combinations

374   and those found in human language.

375

376   **Ethical statement**

377   Our work was carried out under permission from the Limpopo Department of Economic

378   Development, Environment and Tourism (permit number: 001-CPM403-00013) and the Ethical

379   Committee of Pretoria University, South Africa (permit number: EC049-16).

380

381

382   **Competing interests**

383   We have no competing interests.

384

385   **Author contributions**

386   Conceptualization, K.C., S.S., M.B.M., B.B., and S.W.T.; Methodology, K.C., M.B.M., A.N.R., B.B.

387   and S.W.T; Investigation, K.C.; Formal Analysis, K.C., B.B., S.S., S.K.W. and S.W.T; Writing, K.C.,

388   A.N.R., S.S., M.B.M., B.B. and S.W.T.

389

390   **Acknowledgements**

406

407

408

409

410 **References**

411 1. Bousquet CAH, Sumpter DJT, Manser MB. 2010 Moving calls: A vocal mechanism
412   underlying quorum decisions in cohesive groups. *Proc. R. Soc. B Biol. Sci.* **278**, 1482–
413   1488. (doi:10.1098/rspb.2010.1739)

414 2. Blumstein DT. 1995 Golden marmot alarm calls: I. The production of situationally—
415   specific vocalizations. *Ethology* **100**, 113–125.

416 3. Tomasello M. 1999 *The Cultural Origins of Human Cognition Cambridge MA: Harvard*
417   *Univ*. Press.

418 4. Byers BE, Kroodsma DE. 2009 Female mate choice and songbird song repertoires. *Anim.*
419   *Behav.* **77**, 13–22.

420 5. Pika S. 2017 Unpeeling the layers of communicative complexity. *Anim. Behav.* **134**, 223–
421   227.

422 6. Freeberg TM, Dunbar RIM, Ord TJ. 2012 Social complexity as a proximate and ultimate
423   factor in communicative complexity. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **367**, 1785–
424   1801. (doi:10.1098/rstb.2011.0213)

425 7. Pellegrino F, Coupé C, Marsico E. 2011 Across-language perspective on speech
426   information rate. *Language* **87**, 539–558.

427 8. Jackendoff R. 1999 Possible stages in the evolution of the language capacity. *Trends*
428   *Cogn. Sci.* **3**, 272–279. (doi:10.1016/S1364-6613(99)01333-9)

429 9. Nowak MA. 2000 Evolutionary biology of language. *Philos. Trans. R. Soc. B Biol. Sci.*
430   **355**, 1615–1622. (doi:10.1098/rstb.2000.0723)

431 10. Nowak MA, Krakauer DC, Dress A. 1999 An error limit for the evolution of language.
432   *Proc. R. Soc. Lond. B Biol. Sci.* **266**, 2131–2136. (doi:10.1098/rspb.1999.0898)

433 11. Nowak MA, Krakauer DC. 1999 The evolution of language. *Proc. Natl. Acad. Sci.* **96**,
434   8028–8033. (doi:10.1073/pnas.96.14.8028)

435 12. Cäsar C, Zuberbuhler K, Young RJ, Byrne RW. 2013 Titi monkey call sequences vary
436   with predator location and type. *Biol. Lett.* **9**, 20130535–20130535.
437   (doi:10.1098/rsbl.2013.0535)

438 13. Cäsar C, Byrne RW, Young RJ, Zuberbühler K. 2012 The alarm call system of wild
439   black-fronted titi monkeys, Callicebus nigrifrons. *Behav. Ecol. Sociobiol.* **66**, 653–667.
440   (doi:10.1007/s00265-011-1313-0)

441 14. Spillmann B, Dunkel LP, van Noordwijk MA, Amda RNA, Lameira AR, Wich SA, van
442   Schaik CP. 2010 Acoustic properties of long calls given by flanged male orang-utans
443   (Pongo pygmaeus wurmbii) reflect both individual identity and context. *Ethology* **116**,
444   385–395. (doi:10.1111/j.1439-0310.2010.01744.x)

445 15. Arnold K, Zuberbühler K. 2006 Language evolution: Semantic combinations in primate
446   calls. *Nature* **441**, 303–303. (doi:10.1038/441303a)

447 16. Arnold K, Zuberbühler K. 2012 Call combinations in monkeys: Compositional or
448   idiomatic expressions? *Brain Lang.* **120**, 303–309.

449 17. Clarke E, Reichard UH, Zuberbühler K. 2006 The syntax and meaning of wild gibbon
450   songs. *PLoS One* **1**, e73. (doi:0.1371/ journal.pone.0000073)

451 18. Crockford C, Boesch C. 2005 Call combinations in wild chimpanzees. *Behaviour* **142**,
452   397–421. (doi:10.1163/1568539054012047)

453 19. Clay Z, Zuberbühler K. 2009 Food-associated calling sequences in bonobos. *Anim.*
454   *Behav.* **77**, 1387–1396. (doi:10.1016/j.anbehav.2009.02.016)

455 20. Clay Z, Zuberbühler K. 2011 Bonobos extract meaning from call sequences. *PLoS ONE*
456   **6**, e18786. (doi:10.1371/journal.pone.0018786)

457 21. Bouchet H, Pellier A-S, Blois-Heulin C, Lemasson A. 2010 Sex differences in the vocal
458   repertoire of adult red-capped mangabeys (Cercocebus torquatus): a multi-level acoustic
459   analysis. *Am. J. Primatol.* **72**, 360–375. (doi:10.1002/ajp.20791)

460 22. Candiotti A, Zuberbühler K, Lemasson A. 2012 Context-related call combinations in
461   female Diana monkeys. *Anim. Cogn.* **15**, 327–339. (doi:10.1007/s10071-011-0456-8)

462   23. Hurford JR. 2011 *The origins of grammar: Language in the light of evolution II*. Oxford,
463   UK: Oxford University Press.

464   24. Suzuki TN, Zuberbühler K. 2019 Animal syntax. *Curr. Biol.* **29**, R669–R671.
465   (doi:10.1016/j.cub.2019.05.045)

466   25. Ouattara K, Lemasson A, Zuberbühler K. 2009 Campbell's monkeys concatenate
467   vocalizations into context-specific call sequences. *Proc. Natl. Acad. Sci.* **106**, 22026–
468   22031.

469   26. Ouattara K, Lemasson A, Zuberbühler K. 2009 Campbell's monkeys use affixation to
470   alter call meaning. *PloS One* **4**, e7808. (doi:10.1371/ journal.pone.0007808)

471   27. Collier K, Bickel B, van Schaik CP, Manser MB, Townsend SW. 2014 Language
472   evolution: syntax before phonology? *Proc. R. Soc. B Biol. Sci.* **281**.
473   (doi:10.1098/rspb.2014.0263)

474   28. Coye C, Ouattara K, Zuberbuhler K, Lemasson A. 2015 Suffixation influences receivers'
475   behaviour in non-human primates. *Proc. R. Soc. B Biol. Sci.* **282**, 20150265–20150265.
476   (doi:10.1098/rspb.2015.0265)

477   29. Schlenker P *et al.* 2016 Formal monkey linguistics. *Theor. Linguist.* **42**, 1–90.

478   30. Engesser S, Ridley AR, Townsend SW. 2016 Meaningful call combinations and
479   compositional processing in the southern pied babbler. *Proc. Natl. Acad. Sci.* ,
480   201600970.

481   31. Suzuki TN, Wheatcroft D, Griesser M. 2016 Experimental evidence for compositional
482   syntax in bird calls. *Nat. Commun.* **7**, 10986. (doi:10.1038/ncomms10986)

483   32. Jansen DAWAM, Cant MA, Manser MB. 2012 Segmental concatenation of individual
484   signatures and context cues in banded mongoose (Mungos mungo) close calls. *BMC Biol.*
485   **10**, 97. (doi:10.1186/1741-7007-10-97)

486   33. Manser MB, Jansen DAWAM, Graw B, Hollén LI, Bousquet CAH, Furrer RD, le Roux
487   A. 2014 Vocal complexity in meerkats and other mongoose species. In *Advances in the*
488   *Study of Behavior* (eds M Naguib, L Barrett, HJ Brockmann, S Healy, JC Mitani, TJ
489   Roper, LW Simmons), pp. 281–310. Elsevier.

490   34. Déaux EC, Allen AP, Clarke JA, Charrier I. 2016 Concatenation of 'alert' and 'identity'
491   segments in dingoes' alarm calls. *Sci. Rep.* **6**, 30556. (doi:10.1038/srep30556)

492   35. Nowak MA, Plotkin JB, Jansen VAA. 2000 The evolution of syntactic communication.
493   *Nature* **404**, 495–498. (doi:10.1038/35006635)

494   36. Rasa OAE. 1977 The Ethology and Sociology of the Dwarf Mongoose (Helogale
495   undulata rufula). *Z. Für Tierpsychol.* **43**, 337–406. (doi:10.1111/j.1439-
496   0310.1977.tb00487.x)

497   37. Keane B, Waser PM, Creel SR, Creel NM, Elliott LF, Minchella DJ. 1994 Subordinate
498   reproduction in dwarf mongooses. *Anim. Behav.* **47**, 65–75.

499 38. Rasa OAE. 1986 Coordinated vigilance in dwarf mongoose family groups: The
500   'watchman's song' hypothesis and the costs of guarding. *Ethology* **71**, 340–344.
501   (doi:10.1111/j.1439-0310.1986.tb00598.x)

502 39. Collier K, Radford AN, Townsend SW, Manser MB. 2017 Wild dwarf mongooses
503   produce general alert and predator-specific alarm calls. *Behav. Ecol.*
504   (doi:10.1093/beheco/arx091)

505 40. Fichtel C, Perry S, Gros-Louis J. 2005 Alarm calls of white-faced capuchin monkeys: an
506   acoustic analysis. *Anim. Behav.* **70**, 165–176. (doi:10.1016/j.anbehav.2004.09.020)

507 41. Kern JM, Radford AN. 2013 Call of duty? Variation in use of the watchman's song by
508   sentinel dwarf mongooses, Helogale parvula. *Anim. Behav.* **85**, 967–975.
509   (doi:10.1016/j.anbehav.2013.02.020)

510 42. Lachlan RF. 2007 *Luscinia: a bioacoustics analysis computer program*. See
511   luscinia.sourceforge.net.

512 43. Byrt T, Bishop J, Carlin JB. 1993 Bias, prevalence and kappa. *J. Clin. Epidemiol.* **46**,
513   423–429. (doi:10.1016/0895-4356(93)90018-v)

514 44. Sim J, Wright CC. 2005 The Kappa Statistic in Reliability Studies: Use, Interpretation,
515   and Sample Size Requirements. *Phys. Ther.* **85**, 257–268. (doi:10.1093/ptj/85.3.257)

516 45. Belsley DA, Kuh E, Welsch RE. 2005 *Regression diagnostics: Identifying influential
517   data and sources of collinearity*. New York, NY, USA: John Wiley & Sons.

518 46. Mundry R, Sommer C. 2007 Discriminant function analysis with nonindependent data:
519   Consequences and an alternative. *Anim. Behav.* **74**, 965–976.
520   (doi:10.1016/j.anbehav.2006.12.028)

521 47. R Core Team. 2015 *R: A Language and Environment for Statistical Computing*. Vienna,
522   Austria: R Foundation for Statistical Computing. See http://www.R-project.org/.

523 48. Naimi B. 2013 *usdm package*.

524 49. Venables WN, Ripley BD. 2002 *Modern Applied Statistics with S*. 4th edn. New York:
525   Springer.

526 50. Christensen RHB. 2015 *Ordinal - Regression Models for Ordinal Data*. See
527   https://cran.r-project.org/web/packages/ordinal/index.html.

528 51. Bates D, Meachler M, Bolker BM, Walker S. 2015 *lme4: Linear mixed-effects models
529   using Eigen and S4*. See https://cran.r-project.org/web/packages/lme4/index.html.

530 52. Flash T, Hochner B. 2005 Motor primitives in vertebrates and invertebrates. *Curr. Opin.
531   Neurobiol.* **15**, 660–666.

532 53. Lee S-Y, Song X-Y. 2004 Evaluation of the Bayesian and maximum likelihood
533   approaches in analyzing structural equation models with small sample sizes. *Multivar.
534   Behav. Res.* **39**, 653–686.

535   54. Haspelmath M. 2007 Coordination. In *Language typology and syntactic description* (ed T
536   Shopen), pp. 1–51. Cambridge, UK: Cambridge University Press.

537   55. Bradbury JW, Vehrencamp SL. 1998 *Principles of animal communication*. Sunderland,
538   MA, U.S.A.: Sinauer Associates.

539

540

541 Figure 1: Spectrograms of dwarf mongoose alarm calls. 1) Aerial alarm call; 2) Terrestrial alarm call;

542 3) Type 3 or T_3 alarm call composed of two parts: 3.1) pulsed first segment ($T_{3.1}$) and 3.2) noisy

543 second segment ($T_{3.2}$). Window length = 0.05 s, dynamic range = 70 dB.

544

545

546 Figure 2: Frequency of the different types of reaction by focal individuals to the playbacks of natural

547 alarm calls. *: $P<0.05$ [39]; NS: non-significant.

548

549

550 Figure 3: Pairwise comparisons of the relative frequency of different types of reaction by focal

551 individuals in response to playbacks of natural alarm calls and their artificial counterparts. CLMMs:

552 T_3 and T_{3art}: $\chi^2_1=0.22$, p=0.26; aerial and $T_{3.1}$: $\chi^2_1=3.06$, p=0.08; terrestrial and $T_{3.2}$: $\chi^2_1=1.25$, p=0.26.

553 NS: non-significant.

Table 1: Measured acoustic parameters and their definitions. Values presented are mean ± SD. *Used for the pDFA comparing the acoustic structure of aerial, terrestrial and T_3 calls. # used to compare aerial and $T_{3.1}$ calls. ¤ used to compare terrestrial and $T_{3.2}$ calls.

Acoustic parameter	Description	Aerial	Terrestrial	T3	T3.1	T3.2
Call duration (ms) ¤	Time elapsed between the beginning and the end of the call.	374±200	598±169	912±263	370±155	498±198
Overall peak frequency (Hz)*, #, ¤	Peak frequency is the frequency of maximum amplitude within one spectrum of the spectrogram. Overall peak frequency is the frequency of maximum amplitude within the call.	6086±1007	4284±1066	5326±789	5660±861	3891±524
Mean peak frequency (Hz)	Mean of all peak frequencies within the call.	5837±870	3631±527	4315±496	5483±749	3835±400
Maximum peak frequency (Hz)	Peak frequency of highest peak frequency within the call.	7641±1020	5891±1098	7146±1007	7116±1002	4651±722
Minimum peak frequency (Hz)	Peak frequency of the lowest peak frequency within the call.	3526±721	2456±619	3095±481	3567±630	3132±429
Peak frequency start (Hz) *, ¤	Peak frequency at the beginning of the call	4203±704	4141±1094	3850±542	4226±719	3717±530

Peak frequency end (Hz) ¤	Peak frequency at the end of the call	6122±850	3419±522	4399±582	5903±914	3527±394
Mean fundamental frequency (Hz)	Fundamental frequency is the lowest frequency of a periodic waveform.	5838±857	3591±619	4313±495	5495±729	3826±401
Maximum fundamental frequency (Hz)	Fundamental frequency of highest frequency within the call.	7586±993	5830±1136	7113±985	7071±997	4563±733
Minimum fundamental frequency (Hz)	Fundamental frequency of lowest frequency within the call.	3604±749	2434±868	3173±524	3643±632	3238±433
Mean change in peak frequency (Hz)	Mean change in peak frequency over time.	0.63±0.06	0.50±0.01	0.53±0.02	0.61±0.04	0.50±0.01
Change in peak frequency end (Hz) ¤	Change in peak frequency at the end of the call	0.50±0.10	0.51±0.03	0.53±0.02	0.59±0.05	0.49±0.02
Mean change in fundamental frequency (Hz)	Mean change in fundamental frequency over time.	0.63±0.06	0.50±0.01	0.53±0.02	0.61±0.04	0.50±0.01
Minimum change in fundamental frequency (Hz) *	Smallest change in fundamental frequency	0.50±0.11	0.40±0.08	0.46±0.03	0.52±0.07	0.46±0.03
Change in fundamental frequency start (Hz)	Change in the fundamental frequency at the start of the call	0.71±0.04	0.53±0.08	0.55±0.03	0.61±0.05	0.49±0.02

Change in fundamental frequency end (Hz) *, #	Change in the fundamental frequency at the end of the call	0.57±0.10	0.52±0.05	0.52±0.02	0.59±0.05	0.49±0.02
Maximum absolute fundamental frequency change (Hz) #	Highest absolute change in fundamental frequency	0.22±0.06	0.14±0.09	0.19±0.06	0.18±0.06	0.05±0.03
Minimum absolute fundamental frequency change (Hz) ¤	Smallest absolute change in fundamental frequency	0.45±0.06	0.0002±0.0008	$9.90e^{-5}±9.5e^{-5}$	0.05±0.04	0.0002±0.001
Absolute change in fundamental frequency start (Hz) *, ¤	Absolute change in fundamental frequency at the start of the call	0.17±0.06	0.06±0.05	0.06±0.03	0.12±0.04	0.03±0.03
Absolute change in fundamental frequency end (Hz) ¤	Absolute change in fundamental frequency at the end of the call	0.11±0.06	0.03±0.05	0.04±0.02	0.10±0.05	0.01±0.01
Mean Wiener entropy	A measure of noisiness: Ratio of the geometric mean to the arithmetic mean of the power spectrum.	212±3	208±7	215±2	214±2	215±2
Maximum Wiener entropy #, ¤	Highest measure of Wiener entropy	222±1.35	220±2.73	223±0.89	222±1.17	221±1.53

Start Wiener entropy *, #, ¤	Wiener entropy at the beginning of the call	213±3.08	213±6.27	214±3.61	213±3.15	216±3.40
Mean frequency bandwidth (Hz)	Frequency difference between the first and final maximum intensity in the signal.	778±330	1266±688	582±257	621±265	558±213
Maximum frequency bandwidth *, #	Highest frequency bandwidth	2517±851	2400±933	2320±722	2190±729	1279±382
Minimum frequency bandwidth *, #, ¤	Lowest frequency bandwidth	8±60	44±164	0±0	0±0	0±0
End frequency bandwidth *, #, ¤	Frequency bandwidth at the end of the call	250±206	369±588	164±202	228±193	102±208
Number of elements	Number of continuous traces on the spectrogram that compose the call.	4.9±2.4	1.6±0.8	5.9±1.9	4.9±1.9	1.0±0.1
Within syllable gap (ms) *, #, ¤	Total duration of silence between the elements of a call.	202±118	24±38	244±101	197±94	0

Table 2: Simplified contexts in which the different dwarf mongoose alarm calls were produced, both during natural encounters and predator presentations (derived from [39]), as well responses of subjects to the different alarm call types during playback experiments. **X**: call primarily given to this stimulus or main response to the playback of an alarm call. x: alarm call rarely given to a stimulus type or secondary reaction to the playback of an alarm call type. 0: call never given to a stimulus or response never recorded in reaction to the playback of an alarm call

	Production in response to					Main response when hearing alarm call		
	Aerial stimuli	Helium balloon	Dog	Secondary cues	Observer	Run for cover	Vigilance	Look at the sky
Aerial alarm call	**X**	**x**	x	0	**X**	**X**	x	**X**
Terrestrial alarm call	X	0	**X**	**X**	**X**	x	**X**	0
T$_3$ alarm call	**X**	**x**	**X**	x	**X**	**X**	**X**	**X**