Syntax manipulation changes perception of mobbing call sequences across passerine species

Mylène Dutour, Thierry Lengagne, Jean-Paul Léna

1 **Syntax manipulation changes perception of mobbing call**

2 **sequences across passerine species**

3

4 Short running title: **Syntax in birds**

5

6 **Mylène Dutour[1] | Thierry Lengagne[1] | Jean-Paul Léna[1]**

7 [1] Université de Lyon, UMR5023 Ecologie des Hydrosystèmes Naturels et Anthropisés,

8 Université Lyon 1, ENTPE, CNRS, 6 rue Raphaël Dubois, 69622 Villeurbanne, France

9 **Author for correspondence:**

10 Mylène Dutour

11 mylene.dutour@hotmail.com

12 Phone number: +336.72.18.27.62

13

18

19 **CONFLICT OF INTEREST**

20 We have no conflict of interest to declare.

21

22

Abstract

Many species approach predators to harass them and drive them away. Both the intensity of this antipredator strategy and its success are positively related to the size of the group that carries out this mobbing. To recruit individuals to the mob, members of prey species produce mobbing calls. In some songbirds – the Japanese tit, *Parus minor*, and the southern pied babbler, *Turdoides bicolor* – mobbing calls are structurally complex and it has been suggested that they convey information by means of compositional syntax, when meaningful items are combined into larger units. These two species combine alert and recruitment calls into an alert and recruitment sequence when attracting conspecifics to cooperate in mobbing a predator. Whether this rudimentary, two-call, compositional structure is used by other bird species in mobbing calls and how it can alter the ability of heterospecifics to adequately recognize mobbing calls is not well understood. Heterospecifics' responses to mobs are critical to the success of the mobbing strategy, so it is of great importance to understand whether and how syntax influences these responses. To address these questions, we conducted two playback experiments. Firstly, we investigated whether the great tit, *Parus major*, extracts different meanings from different individual motifs (i.e., component calls), and from combined motifs in both natural and artificially-reversed order. We found that great tits extract different meanings from the two motifs involved in mobbing calls, and that they also discriminate for motif order reversal in the mobbing calls sequence. Secondly, we investigated whether heterospecifics (the coal tit, *Periparus ater*, and the common chaffinch, *Fringilla coelebs*) are sensitive to syntax alteration of great tit mobbing calls. While chaffinches did not respond to great tit mobbing calls, coal tits were sensitive to mobbing calls sequence reversal although they did not react in the same way as conspecific subjects. Overall, whereas our results indicate that tits are sensitive to call reversal, this is not to say that tits actually use compositional syntax to increase the information content.

1 | Introduction

During the last decades, accumulating evidence has revealed that animal vocalizations share several features with human language (Collier et al., 2014). Duality of patterning, otherwise known as double articulation (Martinet, 1949), is a property of human language that makes possible a combinatorial structure on two levels: (i) phonological syntax, when meaningless sounds called phonemes (syllable or note) are combined to form meaningful acoustic structures called morphemes (motif) and words; and (ii) compositional syntax, which is the combination of meaningful motifs into a larger structure, whose meaning depends on the motifs involved and the syntactical rules used to put them together (Berwick et al., 2013; Marler, 1998; ten Cate & Okanoya, 2012). The involvement of this property in the acoustic signal enables much more information to be conveyed using a finite set of vocal elements (Berwick et al., 2013). In animals, vocalizations involve a far less complex level of organization than human language. Although more and more evidence for syntax is becoming available for some birds and mammals (Coye et al., 2005; Ouattara et al., 2009), the ability of species to use compositional syntax (or compositionality; see Suzuki et al., 2019) is still debated (Bolhuis et al., 2018; Petkov & Jarvis, 2012; Petkov & Wilson, 2012; Suzuki et al., 2018). More recently, three studies have examined this issue by studying mobbing calls produced by the Japanese tit, *Parus minor*, and the southern pied babbler, *Turdoides bicolor* (Engesser et al., 2016; Suzuki et al., 2016; Suzuki et al., 2017). Mobbing calls are particular forms of alarm signals. They are widespread, especially in passerine birds (Klump & Shalter, 1984). They are emitted by animals trying to chase away a predator (Pettifor, 1990). To be efficient, these calls often share particular features allowing listeners to join the mob, features that are thought to be involved in interspecific communication (Dutour et al., 2017a; Hurd 1996; Marler 1955). In this situation, where calling birds and receivers have a certain level of shared interest, and communication should be clear and as detailed as required to bring about

76 an appropriate response that involves multiple joint behaviours (Cunha et al., 2017), Griesser

77 et al. (2018) have recently suggested from findings from two avian species that compositional

78 syntax may evolve.

79   Mobbing calls have been well described in the Paridae (Carlson et al., 2017a; Jung &

80 Freeberg, 2017) and are usually composed of combinations of frequency modulated elements

81 (Hetrick & Sieving, 2011) referred to hereafter as FME, followed by a string of a repeated

82 loud broadband elements (Templeton et al., 2005), referred to hereafter as D notes. As

83 suggested by several authors, such a pattern could well arise from a hierarchized organization

84 following a 'syntactic-like' rule (Hailman et al., 1985; Hailman & Ficken, 1986; Lucas &

85 Freeberg, 2007). In the Japanese tit, each motif is also used alone in context other than

86 mobbing. The FME motif is used in vigilance situations and it serves as an alarm signal to

87 which receivers respond by scanning for danger (alert calls), while the D notes are involved

88 within foraging flocks and by nest mates to recruit social partners and elicit an approach of

89 the receivers (recruitment calls) (Suzuki et al., 2016; 2017). When the Japanese tits heard the

90 FME motif combined with the D notes as the compound FME-D, which is also a natural

91 vocalization, they showed both scanning and approach behaviours. However, there was little

92 or no response when the call order was artificially reversed to D-FME (i.e. a sequence with

93 unclear or ambiguous meaning; Bolhuis et al., 2018b; Suzuki et al. 2018). From this, authors

94 suggest that compositional syntax is a mechanism for information transmission, although

95 some authors have argued that Suzuki and colleagues do not provide any evidence for genuine

96 compositionality in these bird vocalizations (Bolhuis et al., 2018a; Bolhuis et al., 2018b). One

97 may ask whether the use of a compositional syntax in mobbing calls is specific to *Parus*

98 *minor* or extends to other Paridae species. Furthermore, using compositional syntax to encode

99 information in mobbing calls with several motifs, including ones involved in social cohesion

100 with conspecifics, could limit the ability of heterospecifics to correctly identify mobbing calls.

101 This is particularly intriguing since mobbing calls can communicate the presence of a predator

102 to heterospecifics as well as conspecifics (Dutour et al., 2016; Dutour et al., 2017a; Hurd,

103 1996), prompting the question whether there is actually a similar underlying compositional

104 structure across bird species, enabling them to decode information in heterospecific mobbing

105 calls (Griesser et al., 2018; Russell & Townsend, 2017). It may also be relevant to ask

106 whether species that do not produce combinatorial calls can nevertheless decode them.

107   The main objectives of the present study were first to test whether the great tit, *Parus*

108 *major*, uses compositional syntax in mobbing calls. The great tit is the closest relative of the

109 Japanese tit (Johansson et al., 2013), and these species have similar calls and similar social

110 structures, and also rely on both conspecifics and heterospecifics during mobbing (Randler &

111 Vollmer, 2013; Suzuki, 2016). We investigated whether receivers extract different meanings

112 from FME or D motif alone, and from combined motifs both natural or artificially-reversed

113 order (i.e., FME-D or D-FME sequences). Secondly, we investigated whether heterospecifics

114 are sensitive to syntax alteration of great tit mobbing calls. In this second experiment, we

115 compared the responses of coal tits, *Periparus ater*, and common chaffinches, *Fringilla*

116 *coelebs*, to playbacks of natural and artificially reversed great tit mobbing calls. We chose

117 coal tits rather than other Paridae species since mobbing calls of this species are particularly

118 complex and composed of multiple motifs (Carlson et al., 2017a; Dutour et al., 2017a), and

119 also because our previous work showed that coal tits are especially prone to respond to

120 heterospecific mobbing calls (Dutour et al., 2017a). We selected the chaffinch as an extra-

121 group member (i.e., non-Paridae species) since it is often found in heterospecific mobs

122 although it is less prone than tits to join heterospecific callers (Dutour et al., 2017a). Since the

123 mobbing calls of the chaffinch were composed of a single "chink" note (Randler & Förschler,

124 2011), we predicted that coal tits would be more sensitive to syntax alteration than

125 chaffinches.

126

2 | METHODS

2.1 | Experimental design

Data was collected during playback experiments conducted at the onset of the breeding season

(experiment 1, February/March 2018) and at the end of the breeding season (experiment 2,

July/August 2017) on wild passerines inhabiting mixed deciduous-coniferous forests located

in south-east France (45°80'N, 4°52'E). In order to examine whether great tit mobbing calls

involve compositional syntax (experiment 1), we used five playback types. First, we

examined whether each motif alone (FME calls and D calls) induced a distinct behaviour,

whether vigilance and recruitment. We then examined whether tits hearing the combined

motifs in the natural order (i.e., FME-D calls, the natural mobbing call sequence) display a

combination of the behaviours they exhibit when hearing each motif alone, and we tested

whether this is also the case when the combined motifs are presented in the reverse order (i.e.,

D-FME calls, artificially reversed mobbing call sequence). Finally, we also performed control

tests for which the playback contained only background noise (hereafter referred as BN tests).

We conducted these tests with 100 adult great tits (20 individuals for each test type); each bird

received a single treatment. The goal of experiment 2 was to test whether heterospecific

receivers (coal tits and chaffinches) are sensitive to syntax alteration in the same way as

intraspecific receivers (great tits). To this end, we investigated how individuals of each

species behave when hearing a playback of natural FME-D calls and artificially reversed D-

FME calls of great tits. We also replicated these tests with great tits as receivers in order to

permit comparisons across species. This second experiment involved 90 different individuals

(15 individuals per test and for each species; each individual received a single treatment).

149

2.2 | Field test procedure

151 Field tests were done following a similar methodology used by Dutour et al. (2017) to

152 investigate the response of passerine birds to allopatric mobbing calls. After a focal bird was

153 located, a loudspeaker was placed 30 meter away from the bird at the bottom of a tree. An

154 experimented ornithologist and a field assistant were positioned opposite each other at

155 vantage points at least 15 meters from the loudspeaker and the focal bird to avoid any

156 disturbance during the test. Before the beginning of the experiment, the baseline behaviour of

157 the focal bird was observed during a pre-trial period lasting at least 1 minute. If the bird was

158 found to show alarm behaviour (i.e., emit mobbing calls, which happened in less than 5% of

159 the cases) the test was abandoned. Otherwise, the playback was started when no other

160 passerine was observed near the focal individual. Then, during 1 min of playbacks, two

161 behavioural variables were recorded so as to infer vigilance effort and recruitment propensity,

162 respectively: (1) the number of horizontal scans (we counted the number of obvious

163 movements that birds made with their heads from left to right or right to left, approximately a

164 180 turn; Suzuki et al., 2016) and (2) an approach within a radius of 15 m of the loudspeaker.

165 Horizontal scanning is a good indicator of perceived danger in birds (Curio et al., 1978) and

166 both behaviours, i.e. scanning and approaching, are common during mobbing events (Carlson

167 et al., 2017b; Dutour et al., 2017a; Suzuki et al., 2016; Suzuki et al., 2017). Moreover, we set

168 the approach distance to 15m since this approach distance was previously found to be a

169 relevant measure of mobbing propensity (see Dutour et al., 2017a for more details). All

170 observations were carried out with binoculars. The concordance of the number of scanning

171 between observers was evaluated in a complementary study (see Supporting Information S1).

172 Since this study revealed a very high concordance despite a systemic bias between observers

173 (see Table S1 Supporting Information S1), only the records done by the experienced

174 ornithologist who participated to the whole study were used for data analyses. All trials were

175 conducted between 6 a.m. and 1 p.m. during calm and dry weather. All focal animals were

176   selected to be separated by more than 100 m to minimize pseudo-replication risk. In addition,

177   we never went back twice to the same forest path, and during the breeding season the tits'

178   territories are separated by 50 m in our study area. Once a test was done, we went on more

179   than 100 m before trying to detect another bird which is calling or foraging. No bird was seen

180   following us. Hence, although birds were not individually ringed, the probability of testing the

181   same individual twice was low, and we are confident that our observations were performed on

182   different individuals. Moreover, playback sequences were evenly distributed across the study

183   period to avoid any temporal confounding effect.

184

185   **2.3 | Playback stimuli and playback materials**

186   We used mobbing calls produced by three great tits previously recorded in response to

187   intraspecific mobbing calls (Dutour et al., 2017a). Calls were recorded with a Fostex FR2LE

188   digital recorder connected to a Sennheiser ME67-K6 microphone (see Dutour et al., 2017a for

189   more details). We also used mobbing calls obtained from the Xeno Canto online database

190   (http://www.xeno-canto.org) recorded in different European countries ($n = 6$) located along

191   the species' range in order to encompass the call variation range that a local bird community

192   might experience and to generalize our conclusions. From these recording files, we built 20

193   unique soundtracks of natural mobbing calls (FME-D sequences) using Avisoft-SASLab

194   software (i.e., 20 soundtracks with 1 individual per soundtrack). Recordings were in 16-bit

195   WAV format (44.1 KHz sampling rate). These soundtracks were then used to construct three

196   others, respectively FME calls, D calls, and D-FME calls as follows: FME and D calls were

197   constructed by removing either D or FME calls from each FME-D calls and the D-FME calls

198   were constructed by reversing the order of the motifs in the original FME-D calls of great tits

199   (Figure 1). Within each sound track, calls were repeated at a rate of 26 calls per minute (this

200   calling rate is within the range of the natural repetition rates, unpubl. data). We used a series

201   of five to eight D motifs to construct playback (mean $\pm$ SE = 7.11 $\pm$ 0.06; Figure 1). Each

202   track D notes were placed 82 $\pm$ 40 ms before FME (no difference from the time between FME

203   and D notes in the natural sequence; t = 1.789, $p > 0.05$). The number of D notes could relate

204   to the perceived level of threat (Templeton et al., 2005) which could artificially bring about

205   variation in the response of focal birds. With this mind, all playback sequences were arranged

206   to adjust the ratio of the D notes over the FME notes in a mobbing call sequence, making this

207   ratio slightly lower in our study (range: 5/8 - 8/11) than in Suzuki et al., (2016, range: 7/10 -

208   10/13, see Figure 1 in both manuscripts for comparison). We also constructed 20 control

209   soundtracks using the parts where no birds were calling in the same recordings as natural

210   mobbing calls (BN). In order to avoid pseudoreplication (Kroodsma et al., 2001), we played

211   back each soundtrack only once using a Shopinnov 20 W loudspeaker (the probability of

212   testing an individual twice was low, see above). The average amplitude used for these

213   playbacks was obtained with a sound level meter placed 1 m from the loudspeaker (~ 83 dB,

214   Solo 01dB Metravib, Z weighting, re: 20 μ m Pa).

215

216   **2.4 | Statistical analyses**

217   In the analysis of experiment 1, we used generalized linear mixed models (GLMMs) treating

218   the individual records from which were constructed all subsequent soundtracks (i.e. the test

219   types FME calls, D calls, FME-D calls, D-FME calls and control) as a random effect. The test

220   type was introduced as an explanatory term and the observation duration (i.e., the time during

221   which we could observe the bird) as an adjustment covariate in the fixed part of the model.

222   All analyses performed on the number of scans were performed using a log link function and

223   a generalized Poisson distribution (Proc GLIMMIX; SAS institute inc, 2012) to circumvent

224   the overdispersion of count data relative to a Poisson distribution (see Joe & Zhu, 2005 for

225   justification). For the approaching behaviour, the individual binary response (i.e., approaching

226 versus not approaching) was introduced as the dependent variable using a logit link and a

227 binomial distribution for the error term. For both response variables, the effect of explanatory

228 terms was examined using a non-sequential F test, and subsequently removed from the model

229 if non-significant. Although the random effect was never found significant (results not

230 shown), it was always kept in the model if possible (see below). Multiple comparisons

231 between test types were performed on least square mean estimates using a Dunnett-Hsu

232 correction in the case of pairwise comparisons against the control situation (i.e. BN test) or the

233 Sidak correction otherwise. Finally, for each test type, we used a Wilcoxon test to determine

234 whether scanning behaviour differed between approaching individuals and those staying

235 away, or conversely whether each approach was associated with scanning. Tests were not

236 performed for the FME and the BN playbacks, given the lack of approaching birds in these

237 sequences (see Results).

238   We proceeded in a similar way for the analyses of experiment 2. We first verified that

239 the mobbing call sequence reversal resulted in the same behavioural pattern for conspecific

240 receivers as observed in experiment 1. More specifically, we compared the number of scans

241 and propensity to approach using generalized linear mixed models including the test type and

242 the period (the onset of the breeding season i.e., experiment 1, or the end of the breeding

243 season, i.e., experiment 2) as explanatory terms in the fixed part of the model and the

244 individual from whom the recording came as a random effect.

245   We then compared the responses (scan and approach) of the three species to the great

246 tits mobbing calls presented in the natural order (i.e., FME-D calls) and reversed sequence

247 (i.e., FME-D calls). For this purpose, the receiver species (great tit, coal tit and chaffinch), the

248 test type (FME-D and D-FME calls), and their interactive effect were introduced as

249 explanatory terms in the fixed part of the model and the individual from which the recording

250 came as a random effect. We then performed contrast analyses to test whether each

251   heterospecific receiver specie (i.e. coal tit and chaffinch) responded differently than the

252   conspecific one (i.e. great tit) to the mobbing calls sequence reversion. Multiple comparisons

253   between test types were subsequently performed on least square mean estimates using a Sidak

254   correction. Concerning the approach behaviour, none of the coal tits approached when

255   exposed to one of the two test types (i.e., D-FME calls; $n = 15$) resulting in a sparse data

256   structure. For this reason, likelihood optimization in the presence of a random effect is made

257   impracticable and all tests based on the standard deviation associated to the estimates are

258   made unreliable. We therefore used a GLM instead of a GLMM to analyse the approaching

259   behaviour, and we used likelihood ratio tests instead of the F test to assess the significance of

260   explanatory terms introduced in the model as well as the significance of contrast analyses.

261   Partial analyses for each heterospecific receiver species were then performed, and in the case

262   of the coal tits, the difference between both test types was examined used a Fisher exact test.

263   We also completed these analyses by Wilcoxon tests to determine for each species whether

264   scanning differed between approaching individuals and those staying away or conversely

265   whether each approach was associated with scanning within each test type (i.e., FME-D and D-

266   FME calls). These tests were not performed for coal tits in response to D-FME calls since no

267   individual approached during the test, and nor for chaffinches since almost none of them

268   approached during the test (see Results).

269

270   **2.5 | Ethical note**

271   This work was approved by the Prefecture de l'Ain (DDPP01-15-230) and by the ethical rules

272   set by University Lyon 1 in accordance with the current laws in France.

273

274   **3 | Results**

275   **3.1 | Experiment 1: compositional syntax in great tit mobbing calls**

276   Overall, our analyses indicate that the scanning behaviour varied significantly according to

277   the test type (respectively for the effects of the test type and the observation duration:

278   $F_{4,23} = 14.9, p < 0.0001;\ F_{1,85} = 5.65, p < 0.0197$; see Figure 2a). Pairwise comparisons

279   between each test type and the BN control reveal a significantly higher number of scans for the

280   FME calls and FME-D calls tests (respectively for FME, D, FME-D calls and D-FME: $t = 5.28, p <$

281   $0.0001; t = 0.38, p = 0.98; t = 3.96, p = 0.0021; t = -0.14, p = 0.99$). There was also no

282   significant difference between FME-D calls and FME calls ($t = 1.80, p = 0.59$) and pairwise

283   comparisons confirm that both FME and FME-D calls triggered significantly more scans than D

284   calls and D-FME ones (all $p < 0.01$). The probability of approaching varied significantly

285   between the test types (respectively for the effects of the test type and the observation

286   duration: $F_{4,23} = 4.17, p = 0.011$; $F_{1,85} = 0.94, p = 0.33$; see Figure 2b). Pairwise

287   comparisons between each test type and the BN control reveal a significantly higher approach

288   propensity for the D and FME-D calls tests (respectively for FME, D, FME-D calls and D-FME: $t =$

289   $-0.01, p > 0.99; t = 3.05, p = 0.0154; t = 2.54, p = 0.0464; t = 1.99, p = 0.136$). Moreover,

290   there was no significant difference between D calls and FME-D calls tests ($t = 0.96; p = 0.88$)

291   and pairwise comparisons indicate that both D calls and FME-D calls significantly increased the

292   approach propensity when compared to FME calls but not when compared to D-FME calls (D

293   calls *versus* FME calls : $t = 3.16, p = 0.022$; FME-D calls *versus* FME calls : $t = 2.64, p = 0.071$;

294   D calls *versus* D-FME calls : $t = 1.87, p = 0.32$; FME-D calls *versus* D-FME calls : $t = 0.96, p =$

295   0.88). Finally, regardless of the test type, scanning did not differ between approaching

296   individuals and those staying away (Wilcoxon tests respectively for the D, FME-D, D-FME

297   calls: $W = 49.5, p = 0.78; W = 25.5, p = 0.07; W = 37, p = 0.52$; test not performed for the

298   FME and the BN playbacks given the lack of approaching birds in these sequences).

299

300   **3.2 | Experiment 2: effect of syntax alteration on heterospecific perception**

301   As experiments on heterospecific responses to the syntax alteration of great tit mobbing calls

302   were done at the end of the breeding period while experiment 1 was performed at the onset of

303   the breeding season, we first controlled that great tits response to the alteration of conspecific

304   mobbing calls does not vary during the breeding season. Our results indicate that, whatever

305   the study period (i.e. at the onset or at the end of the breeding season), great tits approached

306   the loudspeaker with the same propensity in response to FME-D calls or D-FME calls (period

307   effect: $F_{1,8} = 0.82; p = 0.39$; test type effect: $F_{1,8} = 1.44; p = 0.26$; interactive effect:

308     $F_{1,8} = 0.01; p = 0.94$; see Figure 3a). However, if the study period also did not alter the

309   effect of the mobbing call inversion on their scanning behaviour, great tits exhibited a higher

310   vigilance effort at the end of the breeding season than at the onset (period effect: $F_{1,8} = $

311     $5.45; p = 0.0478$; test type effect: $F_{1,8} = 32.27; p = 0.0005$; interactive effect: $F_{1,8} = $

312     $0.01; p = 0.93$; see Figure 3b). Moreover, as in the case of experiment 1, scanning did not

313   differ between approaching individuals and those staying away (Wilcoxon test respectively

314   for the FME-D and the D-FME calls: $n = 15$, $W = 29$; $p = 0.86$; $n = 15$, $W = 25$; $p = 0.77$).

315   These results indicate that the intensity of the subjects' reaction to FME-D calls compared to

316   D-FME calls remained unchanged before and after the breeding season for conspecifics, and

317   offers stable ground for interspecific comparison.

318   Comparing the response to the inversion of great tit mobbing call sequence among

319   receiver species, our results reveal a significant interactive effect of the receiver species and

320   the test type on the number of scans displayed during the test (receiver species effect:

321     $F_{2,16} = 2.32; p = 0.131$; test type effect: $F_{1,8} = 9.36; p = 0.0156$; interaction term: $F_{2,16} = $

322     $4.73; p = 0.0243$). As revealed by the contrast analyses, only the difference of the number of

323   scans between the two test types exhibited by chaffinches significantly varied from the one

324   exhibited by the great tits (respectively for chaffinches *versus* great tits and coal tits *versus*

325   great tits : $F_{116} = 7.55; p = 0.0143$; $F_{1,16} = 0.01; p = 0.93$). The difference of approach

326   propensity between the two test types also varied significantly according to the receiver

327   species (receiver species effect: $\chi_2^2 = 20; p < 0.0001$; test type effect: $\chi_1^2 = 2.98; p =$

328   0.084; interaction term: $\chi_2^2 = 6.31; p = 0.042$). However, only the difference of approach

329   propensity between the two test types exhibited by the coal tits significantly varied from the

330   one exhibited by the great tits as revealed by the contrast analyses (respectively for

331   chaffinches *versus* great tits and coal tits *versus* great tits: $\chi_1^2 = 0.82; p = 0.36; \chi_1^2 =$

332    $4.16; p = 0.0413$). Partial analyses indicate that chaffinches did not scan the surroundings

333   differently according to the order of the mobbing call sequence presented ($F_{1,16} = 0.60; p =$

334   0.45). Chaffinches also rarely approached the loudspeaker, regardless of the sequence order

335   (Figure 3c). In contrast, the coal tits displayed both increased vigilance and a higher

336   propensity to approach the loudspeaker when hearing the FME-D calls than when hearing the

337   D-FME calls (for the number of scans: $F_{1,16} = 9.08; p = 0.0082$; for the approach propensity

338   : Fisher exact test $p = 0.042$, see Figure 3b). Finally, unlike the great tits, coal tits

339   approaching the loudspeaker also scanned significantly more intensely (14.2 ± 3.9) than those

340   staying away (8.1 ± 5.2) when hearing the great tit mobbing calls in the natural order

341   (Wilcoxon test: $n = 15$; W = 8; $p = 0.042$; test not performed for the D-FME calls given the

342   lack of approaching birds in this sequence).

343

344   **4 | DISCUSSION**

345   We found that great tits behave distinctively when hearing respectively the FME motifs and

346   the D ones. They scan the environment when hearing the former, and approach the sound

347   source when hearing the latter. As previously seen in the Japanese tits (Suzuki et al., 2016;

348   Suzuki et al., 2017), these results indicate that these two motifs also convey distinct meanings

349   to the great tit. One corresponds to an alarm call (i.e., the FME motif) which elicits increased

350   vigilance from the receiver, while the other corresponds to a recruitment call (i.e., the D

351   motif) which elicits approach behaviour from the receiver. In response to the naturally

352   ordered mobbing call sequence (i.e., FME-D calls), great tits exhibited both increased

353   vigilance (with a high proportion of scanning) and a rapid approach ($n =19$; mean $\pm$ SE =

354   30.22 $\pm$ 3.38 seconds) toward the sound source, potentially to support the simulated caller

355   opposing the putative threat, indicating that tits extracted both meanings from the naturally-

356   ordered combination of the two motifs. To confirm that the mobbing sequence constitutes a

357   combination of two individual calls, a second step is to construct combinations by artificially

358   merging both calls (with the calls originating from the same individual) and show that natural

359   and artificial versions elicit the same response. Both the nature and the intensity of the

360   responses we observed in the present study are similar to those obtained for the Japanese tit

361   (Suzuki et al., 2016) and for the pied babbler (Engesser et al., 2016). However, while in the

362   Japanese tits both the vigilance effort and the propensity to approach vanish when tits hear the

363   mobbing call sequence in the reversed order (Suzuki et al., 2016; Suzuki et al., 2017), the

364   results are different in the case of the great tit. The vigilance effort vanishes but the propensity

365   to approach did not change when the combination order was reversed. Although the

366   percentage of trials in which great tits approached the loudspeaker during the D-FME calls is

367   reduced by 14% compared to the FME-D calls, this difference is low when compared to the

368   one reported by Suzuki et al. (2016) on the Japanese tits (i.e. 41.2% of reduction). Since we

369   observed similar responses in both experiments, which were performed in two breeding

370   seasons, this behavioural pattern is reproducible. Our results therefore suggest that great tits

371   are sensitive to the ordered combination of the two motifs. However, as noted by several

372   authors (Berwick et al., 2013; ten Cate & Okanoya, 2012), this is not to say that tits actually

373   exploit compositional syntax to convey more information. To our knowledge, great tits never

374   use reversed mobbing calls sequences in the wild, and several hypotheses could explain a

375   sequence order reversal effect without invoking compositional syntax usage. First, these

376   results could be explained by the occurrence of a perception bias (Grafe, 1996; Klump &

377   Gerhardt, 1992) when the sequence order is reversed. Indeed, D notes, which are large

378   frequency bandwidths enhancing location, may operate as auditory masking on the FME notes

379   (frequency modulations resistant to degradation; Brown & Handford, 1996; Marler, 1955)

380   given the relative short delay between both sequences. Thus tits could no longer perceive the

381   FME notes when they are artificially placed after the D notes. Such a phenomenon could well

382   explain why great tits reduced vigilance behaviour when hearing the artificially reversed

383   sequence of mobbing calls. This hypothesis could be tested directly by replacing D notes by a

384   distinct high-bandwidth note (e.g., a burst of loud white noise). Auditory masking by D notes

385   could also be the reason why great tits do not produce D-FME calls and, therefore, explain

386   why the ordering rule is the way it is. Mobbing call responsiveness (and/or interpretation)

387   may also depend on the social context, for instance according to the presence of the receiver's

388   mate in the vicinity (Suzuki et al., 2016) or the seasonal activity (Dutour et al., 2017b; Dutour

389   et al., 2019; Lucas et al., 2007). In the Japanese tit, the sequence reversal effect was examined

390   in flock members during the non-breeding season (Suzuki et al., 2016) while in the present

391   study tests were conducted on great tits at the onset and at the end of the breeding season.

392   Since the sensitivity to recruitment calls may vary between these two social contexts, further

393   work is required to examine this point.

394   Chaffinches were barely sensitive to great tit mobbing calls although this species is

395   regularly observed in heterospecific mobbing groups (Dutour et al., 2016). Moreover, we

396   recently found a very low responsiveness of chaffinches to heterospecific mobbing calls when

397   compared to their own mobbing calls (Dutour et al., 2017a; see also Randler & Vollmer,

398   2013). We suggested that such a lack of response could have a causal explanation (the strong

399   dissimilarity between mobbing calls of the chaffinch and other species), and a functional

400   explanation (the exploitation of a distinct ecological niche). The second hypothesis cannot

401 explain our results because chaffinches and great tits usually share the same ecological niche.

402 It is therefore likely that chaffinches are sensitive only to their own species, although the

403 absence of a background noise test in experiment 2 cannot allow us to safely conclude on this

404 point.

405   Conversely, coal tits were not significantly less responsive than conspecific subjects to

406 the naturally-ordered great tit mobbing call sequence (i.e., FME-D calls) regardless of the

407 behaviour being measured (i.e., vigilance or approach). However, their behaviours differ from

408 that of conspecific great tit subjects since approaching coal tits were also more vigilant than

409 those staying away, whereas approaching great tits were not significantly more vigilant than

410 their counterparts staying away. Furthermore, and above all, when the mobbing call sequence

411 order was reversed, coal tits not only reduced their vigilance effort but also no longer

412 approached the loudspeaker. If these results indicate that coal tits are sensitive to great tit

413 mobbing call sequence reversal, they also suggest that coal tits do not extract the same

414 meaning from the mobbing calls sequence as great tits do. In particular, their responsiveness

415 seems more binary or less versatile than that of great tits. This could indicate that some

416 information (i.e., social information) is present in great tit mobbing calls, inducing conspecific

417 responses, whereas coal tits do not pay attention to it. If auditory masking can explain great tit

418 responses to D-FME calls, we hypothesized that wouldn't the same hold true for coal tits. Coal

419 tits could be indeed more sensitive to FME notes when they are artificially placed after the D

420 notes compared to great tits, because their mobbing calls are more complex (Carlson et al.,

421 2017a), composed of combinations of many different calls; and they are more high pitched

422 (unpublished data). One other hypothesis could be used to explain the behaviour of coal tits. It

423 is possible that coal tits acquire the meaning of FME, D and FME-D calls via associative

424 learning (i.e., coal tits learn to associate these calls emitted by great tits with a vigilance

425 situation, a foraging flock or mobbing; e.g., Potvin et al., 2018) because these signals can be

426 heard repeatedly in the field. In contrast, coal tits cannot have acquired any meaning for D-

427 FME sequences as they are never heard. This may explain their lack of reaction to these

428 stimuli. This hypothesis would also be consistent with the 'correlation' between scanning and

429 approach in coal tits (i.e., the binary response depending on whether the birds have been

430 sufficiently exposed to great tits to associate these calls with a danger, or not). Although coal

431 tits are sensitive to great tits mobbing call sequence reversal, our results did not provide

432 evidence for compositional processing. Additional playbacks to verify coal tits' and

433 chaffinches' reactions to FME and D notes, as well as to artificial stimuli starting with an

434 FME- or a D- motif (but finishing with another motif that is normally not used in combination

435 with these motifs) would be most valuable to disentangle this question.

436 Our results also suggest that the combination order of the repetitive loud and broadband

437 notes (i.e., the D motif) preceded by the frequency modulation notes (i.e., the FME motif)

438 within the mobbing call sequence is of prime importance to elicit a response in the receiver. In

439 this respect, it is particularly striking that, in the few bird species for which the effect of

440 compositional syntax has been investigated, all studies reported that natural mobbing call

441 sequences follow this combination ordering (Engesser et al., 2016; Suzuki et al., 2016; Suzuki

442 et al., 2017; Templeton et al., 2005). Furthermore, as suggested by the present study and

443 recent studies on the Japanese tit (Suzuki et al., 2017), it seems that this rule is more

444 important than the acoustic similarity of each motif between species to enable heterospecific

445 response. Nevertheless, whether this rule applies more broadly to bird communities remains

446 to be established and further studies on a larger set of bird species will be necessary. To this

447 aim, comparisons across allopatric species should be especially informative. Furthermore, as

448 birds may not only eavesdrop on heterospecific communication to extract information about

449 predator threats but also about food resources (Magrath et al., 2015), future work is needed to

450   clarify how each motif alone (i.e., alert calls and recruitment calls) is understood by

451   heterospecifics.

REFERENCES

Berwick, R. C., Friederici, A. D., Chomsky, N., & Bolhuis, J. J. (2013). Evolution, brain, and the nature of language. *Trends in cognitive sciences*, *17*, 89–98.

Bolhuis, J. J., Beckers, G. J., Huybregts, M. A., Berwick, R. C., & Everaert, M. B. (2018a). Meaningful syntactic structure in songbird vocalizations?. *PLoS biology*, *16*, e2005157.

Bolhuis, J. J., Beckers, G. J., Huybregts, M. A., Berwick, R. C., & Everaert, M. B. (2018b). The slings and arrows of comparative linguistics. *PLoS biology*, 16(9), e3000019.

Brown, T. J., & Handford, P. (1996). Acoustic signal amplitude patterns: a computer simulation investigation of the acoustic adaptation hypothesis. *Condor*, 608–623.

Carlson, N. V., Healy, S. D., & Templeton, C. N. (2017a). A comparative study of how British tits encode predator threat in their mobbing calls. *Animal Behaviour*, *125*, 77–92.

Carlson, N. V., Pargeter, H. M., & Templeton, C. N. (2017b). Sparrowhawk movement, calling, and presence of dead conspecifics differentially impact blue tit (*Cyanistes caeruleus*) vocal and behavioral mobbing responses. *Behavioral ecology and sociobiology*, *71*, 133.

Collier, K., Bickel, B., van Schaik, C. P., Manser, M. B., & Townsend, S. W. (2014). Language evolution: syntax before phonology?. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20140263.

Cunha, F. C. R. D., Fontenelle, J. C. R., & Griesser, M. (2017). Predation risk drives the expression of mobbing across bird species. *Behavioral Ecology*, *28*, 1517–1523.

Curio, E., Ernst, U., & Vieth, W. (1978). The adaptive significance of avian mobbing: II. Cultural transmission of enemy recognition in blackbirds: effectiveness and some constraints. *Zeitschrift für Tierpsychologie*, *48*, 184–202.

Dutour, M., Lena, J. P., & Lengagne, T. (2016). Mobbing behaviour varies according to predator dangerousness and occurrence. *Animal Behaviour*, *119*, 119–124.

Dutour, M., Léna, J. P., & Lengagne, T. (2017a). Mobbing calls: a signal transcending species boundaries. *Animal Behaviour*, *131*, 3–11.

Dutour, M., Lena, J. P., & Lengagne, T. (2017b). Mobbing behaviour in a passerine community increases with prevalence in predator diet. *Ibis*, *159*, 324–330.

Dutour, M., Cordonnier, M., Léna, J. P., & Lengagne, T. (2019). Seasonal variation in mobbing behaviour of passerine birds. *Journal of Ornithology*, 1–6.

Engesser, S., Ridley, A. R., & Townsend, S. W. (2016). Meaningful call combinations and compositional processing in the southern pied babbler. *Proceedings of the National Academy of Sciences*, *113*, 5976–5981.

Grafe, T. U. (1996). The function of call alternation in the African reed frog (*Hyperolius marmoratus*): precise call timing prevents auditory masking. *Behavioral Ecology and Sociobiology*, *38*(3), 149–158.

Griesser, M., Wheatcroft, D., & Suzuki, T. N. (2018). From bird calls to human language: exploring the evolutionary drivers of compositional syntax. *Current Opinion in Behavioral Sciences*, *21*, 6–12.

Hailman, J. P., Ficken, M. S., & Ficken, R. W. (1985). The 'chick-a-dee'calls of *Parus atricapillus*: a recombinant system of animal communication compared with written English. *Semiotica*, *56*, 191–224.

Hailman, J. P., & Ficken, M. S. (1986). Combinatorial animal communication with computable syntax: Chick-a-dee calling qualifies as" language" by structural linguistics. *Animal Behaviour*.

Hetrick, S. A., & Sieving, K. E. (2011). Antipredator calls of tufted titmice and interspecific transfer of encoded threat information. *Behavioral Ecology*, *23*, 83–92.

Hurd, C. R. (1996). Interspecific attraction to the mobbing calls of black-capped chickadees (*Parus atricapillus*). *Behavioral Ecology and Sociobiology*, *38*, 287–292.

502   Joe, H., & Zhu, R. (2005). Generalized Poisson distribution: the property of mixture of

503   Poisson and comparison with negative binomial distribution. *Biometrical Journal:*

504   *Journal of Mathematical Methods in Biosciences*, *47*, 219–229.

505   Johansson, U. S., Ekman, J., Bowie, R. C., Halvarsson, P., Ohlson, J. I., Price, T. D., &

506   Ericson, P. G. (2013). A complete multilocus species phylogeny of the tits and

507   chickadees (Aves: Paridae). *Molecular phylogenetics and evolution*, *69*(3), 852-860.

508   Jung, H., & Freeberg, T. M. (2017). Variation in chick-a-dee calls of bridled titmice

509   (*Baeolophus wollweberi*): Frequent use of non-combinatorial calls in a combinatorial

510   calling system. *Ethology*, *123*, 835–842.

511   Klump, G. M., & Gerhardt, H. C. (1992). Mechanisms and function of call-timing in male-

512   male interactions in frogs. In *Playback and studies of animal communication* (pp. 153-

513   174). Springer, Boston, MA.

514   Klump, G. M., & Shalter, M. D. (1984). Acoustic behaviour of birds and mammals in the

515   predator context. 1. Factors affecting the structure of alarm signals. 2. The functional-

516   significance and evolution of alarm signals. *Zeitschrift für Tierpsychologie, 66*, 189–226.

517   Kroodsma, D. E., Byers, B. E., Goodale, E., Johnson, S., & Liu, W. C. (2001).

518   Pseudoreplication in playback experiments, revisited a decade later. *Animal Behaviour*,

519   *61*, 1029–1033.

520   Lucas, J. R., & Freeberg, T. M. (2007). *Information and the chick-a-dee call: communicating*

521   *with a complex vocal system. Ecology and behavior of chickadees and titmice.* pp. 199–

522   213. Oxford, Oxford University Press.

523   Lucas, J. R., Freeberg, T. M., Long, G. R., & Krishnan, A. (2007). Seasonal variation in avian

524   auditory evoked responses to tones: a comparative analysis of Carolina chickadees, tufted

525   titmice, and white-breasted nuthatches. *Journal of Comparative Physiology A*, *193*, 201–

526   215.

527 Magrath, R. D., Haff, T. M., Fallow, P. M., & Radford, A. N. (2015). Eavesdropping on

528   heterospecific alarm calls: from mechanisms to consequences. *Biological Reviews*, *90*,

529   560–586.

530 Marler, P. (1955). Characteristics of some animal calls. *Nature*, *176*, 6–8.

531 Marler, P. (1998). *The Origin and Diversification of Language*. (eds Jablonski NG, Aiello

532   LC), pp. 1–20. California University Press.

533 Martinet, A. (1949). La double articulation linguistique. *Travaux du Cercle linguistique de*

534   *Copenhague*, *5*, 30–37.

535 Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's monkeys use affixation to

536   alter call meaning. *PloS one*, *4*, e7808.

537 Petkov, C. I., & Wilson, B. (2012). On the pursuit of the brain network for proto-syntactic

538   learning in non-human primates: conceptual issues and neurobiological hypotheses. *Phil.*

539   *Trans. R. Soc. B*, *367*, 2077–2088.

540 Petkov, C. I., & Jarvis, E. (2012). Birds, primates, and spoken language origins: behavioral

541   phenotypes and neurobiological substrates. *Frontiers in evolutionary neuroscience*, *4*, 12.

542 Potvin, D. A., Ratnayake, C. P., Radford, A. N., & Magrath, R. D. (2018). Birds learn socially

543   to recognize heterospecific alarm calls by acoustic association. *Current Biology*, *28*,

544   2632–2637.

545 Pettifor, R. A. (1990). The effects of avian mobbing on a potential predator, the European

546   kestrel, *Falco tinnunculus*. *Animal Behaviour*, *39*, 821–827.

547 Randler, C., & Förschler, M. I. (2011). Heterospecifics do not respond to subtle differences in

548   chaffinch mobbing calls: message is encoded in number of elements. *Animal behaviour*,

549   *82*, 725–730.

550 Randler, C., & Vollmer, C. (2013). Asymmetries in commitment in an avian communication

551   network. *Naturwissenschaften*, *100*, 199–203.

Russell, A. F., & Townsend, S. W. (2017). Communication: Animal Steps on the Road to Syntax?. *Current Biology*, *27*, R753–R755.

SAS Institute Inc. (2012). SAS/STAT® 12.1 User's Guide. Cary, NC: SAS Institute Inc.

Stoffel, M. A., Nakagawa, S., & Schielzeth, H. (2017). rptR: Repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *8*(11), 1639–1644.

Suzuki, T. N. (2016). Referential calls coordinate multi-species mobbing in a forest bird community. *Journal of Ethology*, *34*, 79–84.

Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature communications*, *7*, 10986.

Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2017). Wild birds use an ordering rule to decode novel call sequences. *Current Biology*, *27*, 2331–2336.

Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2018). Call combinations in birds and the evolution of compositional syntax. *PLoS biology*, *16*, e2006532.

Suzuki, T. N., Griesser, M., & Wheatcroft, D. (2019). Syntactic rules in avian vocal sequences as a window into the evolution of compositionality. *Animal Behaviour*.

Templeton, C. N., Greene, E., & Davis, K. (2005). Allometry of alarm calls: black-capped chickadees encode information about predator size. *Science*, *308*, 1934–1937.

ten Cate, C., & Okanoya, K. (2012). Revisiting the syntactic abilities of non-human animals: natural vocalizations and artificial grammar learning. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *367*, 1984–1994.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

FIGURE CAPTIONS

578 **FIGURE 1** Spectrograms of calls played to great tit, coal tit and chaffinch: **a** FME-D

579 mobbing call of the great tit showing call with introductory frequency modulated elements

580 (FME) (similar to chickadee A or B elements) and subsequent D notes (D) and **b** D-FME call

581 is a reversed combination of FME and D calls. These calls were produced with Avisoft

582 SASLab©

583

584 **FIGURE 2** Responses of great tits to playbacks of FME, FME–D, D–FME, D and control

585 (BN) test types. **a** Number of scans made by tits during playback (generalized linear mixed

586 model: $F_{4,23} = 14.9, p < 0.0001$). Horizontal line: median value; box ends: upper and lower

587 quartiles; whiskers: variation range of values. **b** Percentage of trials in which tits approached

588 the loudspeaker (generalized linear mixed model: $F_{4,23} = 4.17, p = 0.011$). Sample size: $n =$

589 100 individuals. Each individual was exposed to only one per test type, giving $n = 20$ per test

590 type

591

592 **FIGURE 3** Number of horizontal scans and percentage of trials in which **a** great tits, **b** coal

593 tits and **c** chaffinches approached the loudspeaker during the presentations of the playbacks of

594 FME–D and artificially reversed D–FME motifs of mobbing signals. Horizontal line: median

595 value; box ends: upper and lower quartiles; whiskers: variation range of values. Sample size:

596 $n = 90$ individuals ($n = 30$ per species and 15 per test type)