Biological conclusions about importance of order in mobbing calls vary with the reproductive context in Great Tits (Parus major)

Ambre Salis, Thierry Lengagne, Jean-Paul Léna, Mylène Dutour

2   **Biological conclusions about importance of order in mobbing calls vary**

3   **with the reproductive context in Great Tits *Parus major***

4   AMBRE SALIS,[*] THIERRY LENGAGNE, JEAN-PAUL LENA & MYLENE DUTOUR

5   *Univ Lyon, Université Claude Bernard Lyon 1, CNRS, ENTPE, UMR 5023 LEHNA, F-*

6   *69622, Villeurbanne, France.*

7

8   *Corresponding author**.

9   Email: ambre.salis@univ-lyon1.fr

10

11   Currently, there is considerable debate surrounding the presence of some human language

12   specific characteristics in non-human animals, such as the use of compositional syntax (i.e.,

13   meaning of a sequence determined both by meaning of its individual parts and in the way they

14   are combined). Compositional syntax has been investigated in mobbing calls of two closely

15   related tit species, the Japanese Tit *Parus minor* and the Great Tit *Parus major*, but with one

16   contrasting result: hearing calls in the reversed order diminished the behavioural responses of

17   Japanese Tits, however only partially those of Great Tits. This difference may have been due

18   to an external factor such as the season in which the experiment was undertaken, as the Japanese

19   Tits were tested in winter and Great Tits in spring. Here, we studied the responses of Great Tits

20   towards natural and reversed mobbing sequences during spring and winter by investigating two

21   behaviours: approaching and vigilance behaviours. We found that sensitivity to syntax reversal

22   was impacted by the season. The birds were vigilant but less likely to approach reversed calls

23   in winter. However the opposite occurred in spring, with the birds scanning less but still

24   approaching. This study suggests that the perception of combinatorial calls in Great Tits is

25   influenced by the season, emphasizing the importance of context in studies investigating

26   complex cognitive processing in animals.

29   Animal communication is one of the most prolific subjects in animal behaviour studies

30   (Bradbury & Vehrencamp 2011), with one major question being the extent to which we can

31   compare it to human language. Some defining characteristics of human language are

32   classically presented as intentionality, referentiality, vocal learning, and syntax (Hauser *et al*.

33   2002). Numerous comparative studies conducted in animal systems have focused on the first

34   three criteria - intentionality (Seyfarth & Cheney 2003, Graham *et al*. 2019), referentiality

35   (e.g., in Siberian Jays *Perisoreus infaustus*, Griesser 2008), and vocal learning (e.g*.,* song in

36   oscines, Wilbrecht & Nottebohm 2003, Tyack 2019) - and have revealed that these abilities

37   are present in diverse animal species, emphasizing the importance of studying animal

38   communication in our understanding of the evolution of language (Fishbein *et al*. 2019).

39   Indeed, finding such characteristics in our closest relatives may indicate a gradual emergence

40   of language features (Searcy 2019). In contrast, the presence of analogous traits in distant

41   species suggests convergence: the independent rise of the same evolutionary strategy (Searcy

42   2019).

43   Birds are one great example of distant taxa possessing characteristics similar to human

44   language, especially regarding complex combinatorial rules (i.e., a set of principles by which

45   meaning-bearing units can be combined into well-formed complexes, Zuberbühler 2019a).

46   However, previous findings have focused on sequences lacking direct semantic content, with

47   many combinatorial forms being phonocoding (i.e., sound combinations whose individual

48   parts do not possess proper meaning, Engesser & Townsend 2019). Numerous examples of

49   such combinatorial rules can be found in bird songs. For instance, while information about

50   male quality is extracted from the global song, the individual syllables that make up the song

51   do not carry specific information (Catchpole & Slater 1995).

52   Recently, studies have begun to discover various forms of semantic combinatoriality

53   in animals, with some recent research focusing on compositional syntax (Engesser &

54   Townsend 2019, Suzuki *et al*. 2019b, Zuberbühler 2019b). This is observed when, in

55   opposition to phonocoding, the meaning of the whole depends on the meanings of the

56   component parts, and the way they are organized (Hurford 2011). Cases where animals use

57   such compositional structuring have been reported in birds when mobbing predators

58   (Engesser *et al*. 2016, Suzuki *et al*. 2016, Suzuki *et al*. 2017).

59   Mobbing - when a prey moves towards and harasses a predator to chase it away

60   (Carlson *et al*. 2018) - is a suitable behaviour to study compositional syntax. Indeed, prey

61   produce calls to recruit mobbers from a variety of species (Hurd 1996, Randler & Vollmer

62   2013, Dutour *et al*. 2016). In such risky situations (Curio & Regelmann 1986), where callers

63   and receivers have a level of shared interest, and communication should be clear and

64   unambiguous to co-ordinate a response that combines several behaviours, it has been

65   suggested that compositional syntax could emerge (Griesser *et al*. 2018). Experimental studies

66   in the Japanese Tit *Parus minor* provide the strongest case of compositional syntax to date

67   (Suzuki *et al*. 2016, Suzuki *et al*. 2017, but see also Engesser *et al*. 2016 on Southern Pied

68   Babblers *Turdoides bicolor*). Indeed, Suzuki and colleagues demonstrated that the mobbing

69   sequences of Japanese Tits are made through the concatenation of two types of notes -

70   Frequency modulated elements (FME) and D notes (FME-D calls), respectively triggering a

71   vigilance and an approach behaviour - and that order has importance (Suzuki *et al*. 2016).

72   Debates about these results have been profuse (Bolhuis *et al*. 2018a, 2018b, Suzuki *et al*.

73   2018, Townsend *et al*. 2018), and there is a need for broader data on potential

74   compositionality in birds.

75   Dutour *et al*. (2019b) replicated the same experiment on a French population of Great

76   Tits *Parus major,* during the breeding season. This species is the closest relative of the

77   Japanese Tit (Päckert *et al*. 2005, Johansson *et al*. 2013) and produces similarly structured

78   mobbing calls (FME-D calls, Fig. 1). For both species, the FME motif appears to be used in

79   vigilance situations and serves as an alarm signal to which receivers respond by scanning for

80   danger (Suzuki *et al*. 2016, Dutour *et al*. 2019b). The D notes are produced within foraging

81   flocks and by nest mates to recruit social partners and elicit an approach by receivers

82   (recruitment calls, Suzuki *et al*. 2016, Dutour *et al*. 2019b). Therefore, these acoustically

83   distinct calls fulfil the first characteristic of compositionality (i.e., each different part has

84   meaning when produced in isolation). In addition, Great Tits showed excitement signs (wing

85   flicking and body swinging) only toward the combined FME-D calls, and never toward the

86   isolated parts, indicating a meaning more complex than the simple sum of the isolated parts

87   (Salis *et al*. 2020). The second characteristic of compositional structuring reflects the

88   importance of organization, tested through measuring the Great Tits' response when the

89   sequence is reversed (D-FME, demonstrated in the Japanese Tit, Suzuki *et al*. 2016). Here

90   however, almost as many Great Tits approached the speaker to natural sequence (FME-D)

91   than to reversed sequence (D-FME, Dutour *et al*. 2019b). Therefore, the conclusion on syntax

92   use in Great Tits suggested it may not be as clear-cut as for the Japanese Tit and may globally

93   weaken a generalization of syntax results to other species in the genus.

94   One crucial distinction between the studies of Japanese Tits and Great Tits, however,

95   is that the response of Japanese Tits to the inverted sequence was tested during the non-

96   breeding season in winter flocks (Suzuki *et al*. 2016), while these tests were done in a

97   reproductive context at the onset of the breeding season for the Great Tits (Dutour *et al*.

98   2019b). The behaviour of Parids is known to vary markedly between seasons. When breeding,

99   they show enhanced individual territoriality and aggressive behaviour (Shedd 1983, Krams

100   and Krama 2002, Samplonius 2018). In contrast, they join other individuals and form flocks

101   in winter (Hinde 1952, Marra *et al*. 2015). In addition, Great tits approached more

102   heterospecific calls in winter than in the breeding season (Dutour *et al*. 2019a). The attention

103   paid to the same auditory stimuli therefore seems to be dependent on the reproductive state of

104 the birds. We can hypothesize that increased aggressiveness toward intruders in spring may

105 directly affect the behavioural response recorded during syntax-related studies (i.e., approach

106 and scanning, Suzuki *et al*. 2016). In addition, sensibility to conspecific D notes may increase

107 in spring, as pairs use them frequently, and fledglings use similar notes when calling their

108 parents (M.D. pers.obs). In the same way, an increased response toward heterospecific calls

109 (even towards unknown ones, Dutour *et al*. 2017b) in winter potentially indicates an increased

110 attention towards new calls. As a consequence, different biological conclusions could result

111 from the same experiment done in different parts of the life cycle, even if the general syntax-

112 processing ability of the bird remain the same. On a more global scale, determining whether

113 context can impact conclusions on syntax use in Great Tits could be of value in influencing

114 future studies investigating complex language features in animals.

115   Here, we used field playback experiments to examine whether season has an impact on

116 sensitivity to syntax reversion in Great Tits, allowing us to understand the differences found

117 in the two studies investigating the use of syntax in the family Paridae. Specifically, we

118 recorded the behavioural response of Great tits to either naturally ordered calls, reversed calls,

119 or to a control (background noise) in two seasons: in spring, when territoriality is increased,

120 compared to winter, when individuals join other species in flocks.

121

122 **METHODS**

123 **Study sites and experimental design**

124 Data were collected at the start of the breeding season when individual territories are

125 established (February/ March 2018, territories are established in February and defended until

126 June in the study population) and during the winter when tits often form mixed-species flocks

127 (November 2018/ January 2019, Hinde 1952), in the north of Lyon, France.

128   The ecological context of the two seasons is different. In spring, Great Tits invest

129   mainly in breeding (i.e., defending a territory, building a nest, rearing a brood of nestlings,

130   Cramp *et al*. 1993). The defence of the nest against any intruder is intense and well

131   documented (Hinde 1952, Hollander *et al*. 2008). As winter arrives, Great Tits join small

132   conspecific groups and often heterospecific flocks (Carlson *et al*. 2020). Even if some

133   dominance relationship remains (Oberski & Wilson 1991), flocks can combine, and space is

134   divided into undefended overlapping home ranges rather than discrete territories (Ekman

135   1989). Flocks are thought to be an efficient behavioural response toward predator and

136   foraging pressures being different in winter (Sridhar *et al*. 2009). Indeed, increased food

137   storage to counteract the loss of food supply increases predation risk (increase of body mass

138   being negatively correlated with manoeuvrability, Gosler *et al*. 1995), while a decrease in

139   individual territory defence whilst not breeding allows birds to gather without strong

140   competition (Morse 1970).

141   In order to examine whether the response of Great Tits was dependent upon season,

142   we conducted three playback types. First, we examined whether tits hearing the combined

143   motifs in the natural order (FME-D calls, natural mobbing call sequence) displayed a

144   combination of the behaviours classically monitored in mobbing studies (approach and scan,

145   see below for more details), and we tested whether this was also the case when the combined

146   motifs were presented in the reverse order (D-FME calls, artificially reversed mobbing call

147   sequence). Finally, we also performed control tests for which the playback contained only

148   background noise (hereafter referred as BN tests). We conducted these tests on 120 adult

149   Great Tits (20 individuals for each call sequence, at two different seasons).

150

151 **Field test procedure**

152 Each test was conducted by two field assistants. One was assigned to the soundtrack

153 preparation and playback operation, while the other was assigned to the observation of the

154 focal bird. After the localization of a focal bird, the loudspeaker was placed 30 m away from

155 the bird at the bottom of a tree. Both field assistants were positioned opposite each other at

156 vantage points at least 15 m from the loudspeaker and the focal bird to avoid any perturbation

157 during the test. Prior to playback, the baseline behaviour of the focal bird was observed for at

158 least a one-minute pre-trial period. The playback was started when there was no other

159 passerine observed near the focal individual. If the bird showed alarm behaviour before the

160 test (e.g., gave mobbing calls), or if other birds also responded to the playback, the test was

161 abandoned. The sequence was broadcast using a remotely controlled Shopinnov 20 W

162 loudspeaker with an amplitude of $\sim 80 \pm 3$ dB(A) (measured at 1 m from the loudspeaker

163 using Lutron SL-4001, Taipei, Taiwan; C weighting, slow settings, re: 20 μPa), as this is the

164 natural amplitude of Great Tit mobbing calls (from 77 to 80 dB, Templeton *et al*. 2016). The

165 background noise observed in the field (mean $\pm$ sd $= 46.1 \pm 2.1$ dB(C), $n = 20$ measurements)

166 associated to the spherical spread and excess attenuation of sound energy radiating from a

167 source in a deciduous forest (Lengagne & Slater 2002) allowed us to estimate that the

168 broadcast signal reached background noise level at 60-70 m distance.

169 Subsequently, during one minute of playback treatment, two behavioural variables

170 were recorded, respectively, (1) the vigilance effort with the number of horizontal scans (we

171 counted the number of obvious movements that birds made with their heads from left to right

172 or right to left ($\sim 180°$ turn, Suzuki *et al*. 2016) and (2) approach within 15 m of the

173 loudspeaker (Dutour *et al*. 2017b). Scanning for threats is a stereotyped agitation behaviour

174 (Curio 1975, 1978) commonly measured in studies of mobbing response of tits (e.g., Suzuki

175 *et al*. 2016, Carlson *et al*. 2017, Suzuki *et al*. 2017). Both dense vegetation and the small size

176 of these passerine birds prevented a detailed measurement of scanning behaviour using video-

177 recording. These behaviours were thus directly counted an observer using binoculars and

178 reported on a digital audio recorder. To minimize the measurement errors and to make

179 scanning as objective a behaviour as is possible, we considered an individual to scan only if

180 the head movements of the bird from right to left or left to right were obvious. These methods

181 were unable to completely eliminate the potential for unconscious observer bias and we

182 acknowledge that future studies using similar techniques should ideally employ fully blinded

183 designs.

184   All trials were conducted between 06:00 and 13:00 h during calm and dry weather. On

185 average 10 tests were done each day. Selection of all focal animals was based on location,

186 ensuring successively tested animals were always separated by more than 100 m to minimize

187 pseudoreplication risk and ensure that tested birds had not heard a previous test, given that the

188 active space of a sound emitted at 80 dBA with a 46 dB background noise level was less than

189 100 m (Brenowitz 1982, Lengagne & Slater 2002).

190   In addition, we never returned to the same forest path, the size of the research area was

191 large (i.e. ~ 43 km²), and population density in this area is high (M.D. unpubl. data). No bird

192 was seen following us. Hence, although birds were not individually ringed, the probability of

193 testing the same individual twice was low and we are confident that our observations were

194 performed on different individuals.

195

196 **Playback stimuli and materials**

197 We used mobbing calls produced by three Great Tits, previously recorded in response to

198 intraspecific mobbing calls (Dutour *et al*. 2017b). Calls were recorded with a Fostex FR2LE

199 digital recorder (Tokyo, Japan) connected to a Sennheiser ME67-K6 microphone (Wedemark,

200 Germany; see Dutour *et al*. 2017b for more details). We also used mobbing calls obtained

8

201   from the Xeno Canto online database (www.xeno-canto.org) recorded in different European

202   countries ($N = 7$). By controlling several factors when constructing our playbacks (see below),

203   we homogenised our stimuli and reduced potential variation between different geographical

204   areas. We recorded or selected good quality (sampling rate: 44.1 kHz; sample size: 16-bits)

205   recordings, and all playback files were saved as .wav files.

206   From these recording files, we built 20 unique soundtracks of natural mobbing calls

207   (i.e., 20 soundtracks with one individual per soundtrack) using Avisoft-SASLab software

208   (Avisoft Bioacoustics, Glienicke, Germany, one sample of spectrogram and audio file

209   available in Sup.Mat.). These soundtracks were then used to construct the D-FME (i.e.

210   reversed) calls by reversing the motifs order of the original FME-D calls of Great Tits. The

211   same donor individuals were therefore used for both the natural and reversed call sequences.

212   Within each soundtrack, calls were repeated at a rate of 26 calls per minute (characteristics

213   within the range of the natural repetition rates, Dutour *et al*. 2019b). We used a series of five

214   to eight D motifs to construct playback (mean $\pm$ se = 7.11 $\pm$ 0.06). In each track, D notes were

215   placed 82 $\pm$ 40 ms before FME (no difference with time between FME and D notes in the

216   natural sequence; *t*-test = 1.79, $P > 0.05$). All playback sequences were arranged to control the

217   ratio of the D notes over the FME notes in a mobbing call sequence, as note composition is

218   known to alter Great Tits' response (Kalb & Randler 2019). We also constructed 20 control

219   soundtracks using sections where no birds were calling in the same recordings as natural

220   mobbing calls (BN). In order to avoid pseudoreplication (Kroodsma *et al*. 1989), each

221   soundtrack was played back only once.

222

223   **Statistical analysis**

224   All analyses were carried out in R Studio 1.2.5033 (R core development team 2018)

225   using generalized linear mixed models (GLMM, *glmer* package *lme4*). Because the same

226   soundtracks were used in both seasons, soundtrack ID was included as a random effect. To

227   analyse the number of horizontal scans, we used a Poisson error distribution, and a log-link

228   function as no overdispersion was detected (*overdisp.glmer*, package *RVAideMemoire*). We

229   also included the actual time that the bird had been seen during the one-minute trial as offset

230   (57.48 ± 6.36 sec, mean $\pm$ sd) and discarded three individuals (all in the spring-natural order

231   call sequence) for which the time was under 20 *sec* (i.e., sample size for each treatment is 20,

232   except for the spring-natural call sequences with $n = 17$). Approaching behaviour

233   (approaching vs. not approaching, where 1 = response, 0 = no response) was analysed using a

234   binomial distribution for the error term.

235   For both behavioural variables, the same three steps of analysis were carried out: (1)

236   we first ran a model with the three treatments (BN, Natural, and Reverse) at each season. We

237   then extracted the estimated marginal means and pairwise comparisons with related effect

238   sizes (package *emmeans*, function *emmeans*, and package *multcomp*, function *multcomp::cld*).

239   We only focused on five comparisons: in each season, the BN treatment versus the two other

240   treatments, and the difference between the BN treatment in spring and in winter. We therefore

241   used Bonferroni correction with alpha $= 0.05/5 = 0.01$. (2) Secondly, we discarded the BN

242   tests and ran a model only with Natural and Reversed treatments in order to test the

243   significance of the interaction term (Season* Treatment, *Anova* function, package *car*). This

244   term indicates whether the difference between the two treatments is different between

245   seasons. (3) Finally, we reported the effect sizes of the difference between Natural and

246   Reversed treatments for the two seasons (two comparisons: $\alpha = 0.025$). The pairwise

247   comparisons and the related effect sizes of the differences were assessed using odds ratio

248   (hereafter OR, *odds.ratio*, package *questionr*) for the approach behaviour, and with the

249   exponential effect size for the scanning behaviour (package *emmeans* and *multcompview*).

250   Finally, we also verified that approaching behaviour was not correlated to scanning

251   behaviour using Wilcoxon tests for each season and each playback type. As revealed by these

252   analyses, no association between these two behaviours was detected (all $P > 0.05$).

253

254   **RESULTS**

255   **Scanning Behaviour**

256   In both seasons, Great Tits scanned less toward the control treatment (BN =

257   background noise) than toward the two other call sequences (all $P < 0.001$, Fig. 2, Table 1a),

258   except for the reversed call sequences in spring ($P = 0.89$, Table 1a). There was no significant

259   difference in scan number between seasons for BN soundtracks ($P = 0.03$, $\alpha = 0.01$, Table 1a).

260   The interaction between treatment and season was statistically significant when the

261   control tests were discarded (Table 1b), with a substantially larger difference between

262   acoustic treatments in spring than in winter. Indeed, Great Tits produced 2.29 more scans

263   toward natural calls than toward reversed calls in spring (Table 1c, Fig. 2), while there was no

264   such difference between treatments in winter (ratio of 1.3, Table 1c, Fig. 2). The treatment

265   term was also statistically significant, indicating an overall reduced response to reversed calls

266   ($P < 0.001$, Table 1b).

267

268   **Approaching behaviour**

269   Respectively 50% and 65% of Great Tits approached natural calls in spring and

270   winter. In contrast, only 5% of Great tits in spring, and none of them in winter, approached

271   toward the controls (Fig 3). The difference between natural calls and controls was statistically

272   significant in both seasons (Table 1a). In contrast, no statically significant difference in

273   probability of approach was detected between reversed calls and controls (35% in spring, 25%

274   in winter, Fig 3, Table 1a). There was almost no approach exhibited to control playbacks and

11

275   so the difference between controls in spring and in winter was not significant ($P = 0.49$, Table

276   1a).

277   The interaction between treatment and season was not statistically significant in our

278   model ($P = 0.25$, Fig 3, Table 1b), but the treatment term was statistically significant (more

279   approach to the natural treatment, $P = 0.02$, Fig 3). Nonetheless, even though no interaction

280   was detected, we found a statistically significant difference between reversed and natural

281   playback in winter (OR = 5.57, Table 1c), but not in spring (OR = 1.86), indicating that Great

282   Tits reduced their response toward reversed calls in winter, but approached both treatments

283   similarly in spring.

284

285   **DISCUSSION**

286   Great Tits reduced their overall behavioural response toward reversed calls compared

287   to natural ones. However, season affected the strength of that effect. Great Tits strongly

288   reduced their scanning behaviour when hearing reversed calls compared to natural calls in

289   spring, but such a reduction was not statistically significant in winter. In contrast, for the

290   approach behaviour, the difference between natural and reversed call sequences was

291   statistically significant in winter, but not in spring. The compositional syntax hypothesis

292   requires a decrease in behavioural responses (i.e., scanning and approaching) toward reversed

293   call sequences. Here, we demonstrate that testing birds in different seasons would have

294   resulted in different biological conclusions about syntax-use ability in tits.

295   For the approach behaviour, the difference between natural and reversed calls was

296   significant in winter, but not in spring. Importantly however, the interaction in our model was

297   not statistically significant. This may be due to the response variable being binary, thus

298   limiting power to detect effects with the available sample sizes. Nonetheless, the fact that this

12

299   difference was strong in spring but not in winter is some evidence of an effect of season on

300   the response to different treatments by Great Tits. The absence of variation in the probability

301   of approach in spring is consistent with the results of Dutour *et al*. (2019b) in which the same

302   protocol was used, and could therefore explain the difference with the study of Suzuki *et al*.

303   (2016), in which birds were tested in winter. These results therefore suggest that season does

304   affect Great Tit behavioural responses to natural and reversed playback.

305   Seasonal effects on responses to different syntactic forms has previously been

306   demonstrated in birds. For example, a similar result was found by Clucas *et al*. (2004)

307   investigating the vocal response of Carolina chickadees *Poecile carolinensis* when confronted

308   with familiar (AAADDD) and unfamiliar order calls (ADADAD): birds discriminated the

309   familiar from the unfamiliar calls in winter, but not in summer. Both Carolina Chickadees and

310   Great Tits live in mixed species flocks during winter (Hinde 1952), which are thought to

311   increase both foraging efficiency and defence against predators (Sridhar *et al*. 2009).

312   Vigilance toward heterospecific and potentially unfamiliar calls may therefore increase in

313   winter, as contact with heterospecifics with respect to food and alarm calls is common (Morse

314   1970, Clay *et al*. 2012). Corroborating this hypothesis, Great Tits increased their response

315   toward heterospecifics in winter compared to spring (Dutour *et al*. 2019a). We can therefore

316   hypothesize that in winter, birds are highly vigilant and more efficient in discriminating

317   unfamiliar calls, consequently leading to high scanning behaviour and low approach

318   behaviour toward artificial sequences. In contrast, during the breeding season, the response of

319   Great Tits is probably linked to strong competition and reproductive investment expressed

320   through increased aggression (Wingfield & Ramenofsky 1985). This leads to a more risk

321   prone behaviour: approaching toward any calls similar to recruitment calls (i.e., calls with D

322   notes, Kalb *et al*. 2019) without necessarily scanning. Further research is clearly required to

323   test this hypothesis and to understand why the response to reversed ordering of the mobbing

324 call sequence varies according to the season. In particular, comparisons of the response to the

325 isolated recruitment call (FME notes) and alarm call (D notes) according to the season is still

326 lacking. Such experimental analysis is required to fully appreciate the underlying reason for

327 scanning and approaching behaviours varying according to the season. Moreover, testing such

328 a hypothesis could help understand whether during the breeding season, Great Tits are indeed

329 more sensitive to recruitment calls, irrespective of the syntactic form.

330   Our findings reveal that in spring, we could not conclude that Great Tits reduced their

331 approach behaviour when hearing reversed calls. On the opposite, the reduced scanning

332 response was statistically significant in spring, but not in winter. Consequently, the same

333 experiment done in different seasons would not have led to the same biological conclusions.

334 Indeed, approaching and scanning more toward natural mobbing call sequences than toward

335 reversed call sequences is an essential prerequisite for compositional syntax in birds (Suzuki

336 *et al*. 2016). Concluding whether Great Tits do truly use syntax is consequently difficult: it is

337 unlikely that high mental ability such as syntax use vary with the seasons; but we show that

338 the external proof of it (the bird's behaviour) does. This may reflect a change of priorities in

339 birds between seasons (heterospecific co-operation in winter versus high territorial defence in

340 spring). In fact, such a problem may be common in language-related studies of animal

341 behaviour (Graham *et al*. 2019). For example, deciphering whether chimpanzees *Pan*

342 *troglodytes* possess a zero or second order intentionality level (i.e., when the sender does or

343 does not intend to signal and attribute mental state to the receiver; Townsend *et al*. 2017) is

344 dependent on aspects of predator stimuli (Crockford *et al*. 2012, Schel *et al*. 2013, Graham *et*

345 *al*. 2019).

346   Even if the effect sizes of the differences varied between the two seasons, we found an

347 overall reduced response in both scanning and approaching behaviours in response to the

348 reversed mobbing calls. This is consistent with previous studies of bird syntax (Suzuki *et al*.

14

349 2016, 2017, Dutour *et al*. 2019b). Discussion of whether this reduced response does support

350 the presence of compositional syntax in birds has been intense (see Bolhuis *et al*. 2018a,

351 2018b, Suzuki *et al*. 2018). Indeed, active use of compositional syntax by birds or simpler

352 mechanisms such as the masking hypothesis (D notes put in the first place acoustically hiding

353 the FME notes, Dutour *et al*. 2019b, Dutour *et al*. 2020) are still to be tested. We advocate for

354 the monitoring of other fine behavioural cues such as wing flicking, tail flit or calling in order

355 to provide a more reliable overview of the bird's internal state, (Cully & Ligon 1976, Carlson

356 *et al*. 2017, Salis *et al.* 2020). Other lines of inquiry, for example in neurobiology, may

357 complete field data and allow a broader view of birds' abilities (Marler 2004). Experiments on

358 Black-capped Chickadees *Poecile atricapillus* brains have shown, for example, that the neural

359 response (i.e., protein expression important for processing and categorizing conspecific

360 vocalizations) increased with the level of threat conveyed in the calls (Avey *et al*. 2011).

361 Consequently, going one step further by studying neural responses to different syntactic forms

362 and thus, the neural mechanisms underlying the production and perception of call

363 combinations could be investigated (Suzuki *et al*. 2019a).

364

365   In conclusion, context can affect the behavioural cues used to assess syntax

366 understanding by Great Tits, and thus affect biological conclusions about syntax use in this

367 species. Thorough work on the impact of context on animal behaviour could turn out to be of

368 critical importance and directly impact future studies investigating high cognitive processing

369 in non-human animals.

370

371

372

380

381

382 **REFERENCES**

383 **Avey, M. T., Hoeschele M., Moscicki M.K., Bloomfield L.L. & Sturdy, C.B.** 2011. Neural

384 Correlates of Threat Perception: Neural Equivalence of Conspecific and Heterospecific

385 Mobbing Calls Is Learned. PLoS One **6**: e23844.

386 **Bolhuis, J. J., Beckers G.J.L., Huybregts, M.A.C., Berwick R.C. & Everaert, M.B.H.**

387 2018a. Meaningful syntactic structure in songbird vocalizations? PLoS Biol. **16**: e2005157.

388 **Bolhuis, J. J., Beckers G.J.L., Huybregts, M.A.C., Berwick R.C. & Everaert, M.B.H.**

389 2018b. The slings and arrows of comparative linguistics. PLoS Biol. **16**: e3000019.

390 **Bradbury, J.W. & Vehrencamp, S.L.** 2011. *Principles of animal communication,* 2nd

391 edition*.* Sunderland: Sinauer Associates.

392 **Brenowitz, E. A.** 1982. The active space of red winged blackbird song. J. Comp. Physiol.

393 **147**: 511–522.

394 **Carlson, N. V., Healy S.D. &Templeton, C.N.** 2018. Mobbing. Curr. Biol. **28**:

395 R1081-R1082.

396 **Carlson, N. V., Pargeter H.M. &Templeton, C.N.** 2017. Sparrowhawk movement, calling,

397 and presence of dead conspecifics differentially impact Blue Tit (*Cyanistes caeruleus*) vocal

398 and behavioral mobbing responses. Behav. Ecol. Sociobiol. **71**: 133.

399 **Carlson, N. V., Healy, S. D. & Templeton, C. N.** 2020. What makes a 'community

400 informant'? Reliability and anti-predator signal eavesdropping across mixed-species flocks of

401 tits. Anim. Behav. Cogn. **7**: 214-246.

402 **Catchpole, C. K. & Slater, T. L. B.** 1995. *Bird song: Themes and variations*. Cambridge:

403 Cambridge University Press.

404 **Clay, Z., Smith, C.L. & Blumstein, D.T.** 2012. Food-associated vocalizations in mammals

405 and birds: What do these calls really mean? Anim. Behav. **83**: 323-330.

406 **Clucas, B.A., Freeberg, T.M. & Lucas, J.R.** 2004. Chick-a-Dee Call Syntax, Social

407 Context, and Season Affect Vocal Responses of Carolina Chickadees (*Poecile Carolinensis*).

408 Behav. Ecol. Sociobiol. **55**: 187-96.

409 **Cramp, S., Perrins, C. M. & Brooks, D. J.** 1993. *Handbook of the birds of Europe the*

410 *middle east and north Africa Vol. 7: Flycatchers to shrikes*. Oxford and New York: Oxford

411 University Press.

412 **Crockford, C., Wittig, R.M., Mundry, R. & Zuberbühler, K.** 2012. Wild Chimpanzees

413 Inform Ignorant Group Members of Danger. Curr. Biol. **22**: 142-146.

414 **Cully, J. F. & Ligon, J.D.** 1976. Comparative Mobbing Behavior of Scrub and Mexican

415 Jays. Auk **93**: 116-125.

416 **Cumming, G., Fidler, F. & Vaux, D. L.** 2007. Error bars in experimental biology. J. Cell

417 Biol. **177**: 7-11.

418 **Cumming, G. & Finch, S.** 2005. Inference by Eye: Confidence Intervals and How to Read

419 Pictures of Data. Am. Psychol. **60**: 170-180.

420   **Curio, E.** 1975. The functional organization of anti-predator behaviour in the pied flycatcher:

421   a study of avian visual perception. Anim. Behav, **23**: 1-115.

422   **Curio, E.** 1978. The adaptive significance of avian mobbing: I. Teleonomic hypotheses and

423   predictions. Ethology **48**: 175-183.

424   **Curio, E. & Regelmann, K.** 1986. Predator harassment implies a real deadly risk: A reply to

425   Hennessy. Ethology **72**: 75-78.

426   **Dutour, M., Cordonnier, M., Léna, J.-P. & Lengagne, T.** 2019a. Seasonal variation in

427   mobbing behavior of passerine birds. J. Ornithol. **160**: 509-514.

428   **Dutour, M., Léna, J.-P. & Lengagne, T.** 2016. Mobbing behavior varies according to

429   predator dangerousness and occurrence. Anim. Behav. **119**: 119-124.

430   **Dutour, M., Léna, J.-P. & Lengagne, T.** 2017b. Mobbing calls: a signal transcending

431   species boundaries. Anim. Behav. **131**: 3-11.

432   **Dutour, M., Léna, J.-P. & Lengagne, T.** 2019b. Syntax manipulation changes perception of

433   mobbing call sequences across passerine species. Ethology **125**: 635-644.

434   **Dutour, M., Suzuki, T.N. & Wheatcroft, D.** 2020. Great tit responses to the calls of an

435   unfamiliar species suggest conserved perception of call ordering. Behav. Ecol. Sociobiol. **74**:

436   1-9.

437   **Ekman, J.** 1989. Ecology of Non-Breeding Social Systems of Parus. Wilson Bull. **101**: 263-

438   288.

439   **Engesser, S., Ridley, A.R. & Townsend, S.W.** 2016. Meaningful call combinations and

440   compositional processing in the southern pied babbler. Proc. Natl. Acad. Sci. U.S.A. **113**:

441   5976-5981.

442   **Engesser, S. & Townsend, S.W.** 2019. Combinatoriality in the vocal systems of nonhuman

443   animals. Wires. Dev. Biol. **10**: e1493.

444   **Fishbein, A.R., Fritz, J.B., Idsardi, W.J. & Wilkinson, G.S.** 2019. What can animal

445   communication teach us about human language? Philos. T. Roy. Soc. B **375**: 20190042.

446   **Gosler, A.G., Greenwood, J.J.D & Perrins, C.** 1995. Predation risk and the cost of being

447   fat. Nature **377**: 621–623.

448   **Graham, K. E., Wilke, C., Lahiff, N.J. & Slocombe, K.E.** 2019. Scratching beneath the

449   surface: intentionality in great ape signal production. Philos. T. Roy. Soc. B **375**: 20180403.

450   **Griesser, M.** 2008. Referential Calls Signal Predator Behavior in a Group-Living Bird

451   Species. Curr. Biol. **18**: 69-73.

452   **Griesser, M., Wheatcroft, D. & Suzuki, T.N.** 2018. From bird calls to human language:

453   exploring the evolutionary drivers of compositional syntax. Curr. Opin. Behav. Sci. **21**: 6-12.

454   **Hauser, M. D., Chomsky, N. & Fitch, W.T.** 2002. The Faculty of Language: What Is It,

455   Who Has It, and How Did It Evolve? Science **298**: 1569-79.

456   **Hinde, R. A.** 1952. The Behavior of the Great Tit (*Parus major*) and Some Other Related

457   Species. Behaviour **2**: 1-201.

458   **Hollander, F. A., Van Overveld, T., Tokka, I. & Matthysen, E.** 2008. Personality and Nest

459   Defence in the Great Tit (*Parus major*). Ethology **114**: 405-412.

460   **Hurd, C. R**. 1996. Interspecific attraction to the mobbing calls of Black-Capped Chickadees

461   (*Parus atricapillus*). Behav. Ecol. Sociobiol. **38**: 287-292.

462   **Hurford, J.R.** 2011. *The linguistic theory of nume*rals. Cambridge: Cambridge University

463   Press.

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

465   **Ericson, P. G.** 2013. A complete multilocus species phylogeny of the tits and chickadees

466   (Aves: Paridae). Mol. Phylogenetics Evol. **69**: 852–860.

467   **Kalb, N., Anger, F. & Randler, C.** 2019. Great Tits encode contextual information in their

468   food and mobbing calls. R. Soc. Open Sci. **6**: 191210.

469   **Kalb, N. & Randler, C.** 2019. Behavioral Responses to Conspecific Mobbing Calls Are

470   Predator-specific in Great Tits (*Parus major*). Ecol. Evol. **9**: 9207-13.

471   **Krams, I. & Krama, T.** 2002. Interspecific reciprocity explains mobbing behavior of the

472   breeding chaffinches, *Fringilla coelebs*. Proc. R. Soc. Lond. B. **269**: 2345-2350.

473   **Kroodsma, D. E.** 1989. Suggested experimental designs for song playbacks. Anim. Behav.

474   **37**: 600-609.

475   **Lengagne, T. & Slater, P.J.B.** 2002.The effects of rain on acoustic communication: tawny

476   owls have good reason for calling less in wet weather. Proc. R. Soc. Lond. B. **269**: 2121–

477   2125.

478   **Marler, P.** 2004. Bird Calls: Their Potential for Behavioral Neurobiology. Ann. N. Y. Acad.

479   Sci. **1016**: 31-44.

480   **Marra, P. P., Cohen, E.B., Loss, S.R., Rutter, J.E. & Tonra, C.M.** 2015. A call for full

481   annual cycle research in animal ecology. Biol. Lett. **11**: 20150552.

482   **Morse, D. H.** 1970. Ecological Aspects of Some Mixed-Species Foraging Flocks of Birds.

483   Ecol. Monogr. **40**: 119-168.

484   **Oberski, I. M., & Wilson, J. D**. 1991. Territoriality and site-related dominance: on two

485   related concepts in avian social organization. Ethology **87**: 225-236.

486   **Päckert, M., Martens, J., Eck, S., Nazarenko, A.A., Valchuk, O.P., Petri, B. & Veith, M.**

487   2005. The great tit (Parus major) - a misclassified ring species. Biol. J. Linn. Soc. **86**:

488   153-174.

489   **Randler, C. & Vollmer, C.** 2013. Asymmetries in commitment in an avian communication

490   network. *Sci. Nat.* **100**: 199-203.

491   **R Core Development Team** (2018). R: A Language and Environment for Statistical

492   Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-

493   project.org/

494 **Salis, A., Léna, J. P. & Lengagne, T.** (2020). Great tits (*Parus major*) adequately respond to

495 both allopatric combinatorial mobbing calls and their isolated parts. Ethology **00**:1-10.

496 **Samplonius, J.M.** Interspecific aggression declines seasonally in breeding great tits Parus

497 major. 2018. Ethology **00**:1–5

498 **Schel, A. M., Townsend, S.W., Machanda, Z., Zuberbühler, K. & Slocombe**, K.E. 2013.

499 Chimpanzee Alarm Call Production Meets Key Criteria for Intentionality. PLoS One **8**:

500 e76674.

501 **Searcy, W. A.** 2019. Animal communication, cognition, and the evolution of language. Anim.

502 Behav. **151**: 203-205.

503 **Seyfarth, R. M. & Cheney, D.L.** 2003. Signalers and Receivers in Animal Communication.

504 Annu. Rev. Phychol. **54**: 145-173.

505 **Shedd, D. H.** 1983. Seasonal Variation in Mobbing Intensity in the Black-Capped Chickadee.

506 Wilson J. Ornithol. **95**: 343-348.

507 **Sridhar, H., Beauchamp, G. & Shanker, K.** 2009. Why do birds participate in mixed-

508 species foraging flocks? A large-scale synthesis. Anim. Behav. **78**: 337-347.

509 **Suzuki, T. N., Griesser, M. & Wheatcroft, D.** 2019a. Syntactic rules in avian vocal

510 sequences as a window into the evolution of compositionality. Anim. Behav.**151**: 267-274.

511 **Suzuki, T. N., Wheatcroft, D. & Griesser, M.** 2016. Experimental evidence for

512 compositional syntax in bird calls. Nat. Commun. **7**: 10986.

513 **Suzuki, T. N., Wheatcroft, D. & Griesser, M.** 2017. Wild Birds Use an Ordering Rule to

514 Decode Novel Call Sequences. Curr. Biol. **27**: 2331-2336.e3.

515 **Suzuki, T. N., Wheatcroft, D. & Griesser, M.** 2018. Call combinations in birds and the

516 evolution of compositional syntax. PLoS Biol. **16**: e2006532.

517 **Suzuki, T. N., Wheatcroft, D. & Griesser M.** 2019b. The syntax–semantics interface in

518 animal vocal communication. Philos. T. Roy. Soc. B **375**: 20180405.

519 **Templeton, C.N., Zollinger, S.A. & Brumm, H.** 2016. Traffic noise drowns out Great Tit

520 alarm calls. Curr. Biol., **26**: R1173-R1174.

521 **Townsend, S. W., Koski, S.E., Byrne, R.W., Slocombe, K.E., Bickel, B., Boeckle, M.,**

522 **Braga Goncalves, I., Burkart, J.M., Flower, T., Gaunet F.,** *et al.* 2017. Exorcising Grice's

523 ghost: An empirical approach to studying intentional communication in animals. Biol. Rev.

524 **92**: 1427-1433.

525 **Townsend, S.W., Engesser, S., Stoll, S., Zuberbühler K. & Bickel B.** 2018.

526 Compositionality in animals and humans. PLoS Biol. **16**: e2006425.

527 **Tyack, P. L.** 2019. A taxonomy for vocal learning. Philos. T. Roy. Soc. B **375**: 20180406.

528 **Wilbrecht, L. & Nottebohm F.** 2003. Vocal learning in birds and humans. Ment. Retard.

529 Dev. D. R. **9**: 135-148.

530 **Wingfield J.C. & Ramenofsky M.** 1985. *Testosterone and Aggressive Behaviour during the*

531 *Reproductive Cycle of Male Birds*. In Neurobiology: Current Comparative Approaches. (R.

532 Gilles, J. Balthazart, Editors). Berlin: Springer.

533 **Zuberbühler, K**. 2019a. Evolutionary roads to syntax. Anim. Behav. **151**: 269-265.

534 **Zuberbühler, K.** 2019b. Syntax and compositionality in animal communication. Philos. T.

535 Roy. Soc. B. **375**: 20190062.

536 **SUPPORTING INFORMATION**

537 **Figure 1.** Spectrograms of call treatment played to Great Tits. (a) FME-D calls and (b) D-

538 FME calls, which are a reversed combination of FME and D calls. These calls were digitally

539 edited using AvisoftSASLab Pro software.

540 **Table 1.** Overview of the playback stimuli used for the study. XC is the acronym for xeno-

541 canto.

542

544   **Table 1.** Details of the pairwise comparisons (a, c) and model output (b). (a) Refers to the

545   comparisons with the control (BN), (b) to the model in which control tests were discarded,

546   and (c) to the comparison between natural and reversed calls in the two seasons studied. For

547   scanning behaviour, the ratio (effect size) indicates the strength of the difference between the

548   first and the second term (e.g., Great Tits scanned on average 1.44 times more toward

549   background noise in spring than in winter). The same reasoning applies for the odds ratio

550   (OR) for the approach behaviour. A GLMM with a Poisson error distribution was used for the

551   number of scans (~180° horizontal head turns) during 1min. A GLMM with a binomial error

552   distribution was used for approach behaviour. The pairwise comparisons were extracted from

553   the model and *P*-values were corrected for multiple comparisons (Bonferroni correction: α =

554   0.01 for (a), α = 0.05 for (b), and α = 0.025 for (c) due to correction for multiple

555   comparisons). se, standard error; *z*, z statistic; *P,* p-value (bold= statistically significant); χ^2,

556   Chi-squared test; df, degree of freedom; 95%CI, 95% confidence intervals.

557

|	SCAN				APPROACH			
(a) Differences with control BN	Ratio	se	z	P	OR	95%CI	z	P
BN Winter- BN Spring	1.44	0.24	2.2	0.03	3.15	[0.12; 82.17]	0.69	0.49
BN Winter- Natural Winter	3.13	0.53	6.8	**< 0.001**	73.8	[3.89; 1401.64]	2.86	0.**004**
BN Winter- Reversed Winter	2.41	0.41	5.11	**< 0.001**	14.54	[0.75; 283.38]	1.77	0.08
BN Spring - Natural Spring	2.34	0.36	5.46	**< 0.001**	19	[2.12; 170.39]	2.63	**0.008**
BN Spring- Reversed Spring	1.02	0.17	0.13	0.89	10.23	[1.12; 93.35]	2.06	0.04
(b) Model discarding BN tests	X²	df	P		X²	df	P	
Season (Spring or Winter)	3.52	1	0.06		0.05	1	0.82	
Treatment (Natural or Reversed)	18.26	1	**< 0.001**		5.64	1	**0.02**	
Season* Treatment	11.49	1	**< 0.001**		1.33	1	0.25	
(c) Difference Natural-Reversed calls	Ratio	se	z	P	OR	95%CI	z	P
Natural Winter- Reversed Winter	1.3	0.18	1.92	0.05	5.57	[1.42; 21.86]	2.46	**0.01**
Natural Spring- Reversed Spring	2.29	0.35	5.45	**< 0.001**	1.86	[0.52; 6.61]	0.96	0.33

558

FIGURES

560

561

562

563

564

565
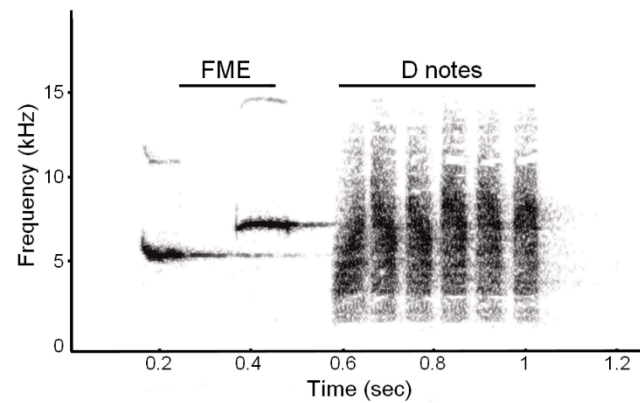
566 **Figure 1.** Sonogram of a Great Tit mobbing call, comprising two distinct calls: (i) alert call

567 (composed of Frequency Modulated Elements, FME) and (ii) recruitment call (a string of

568 notes with broadband frequencies, D notes). This mobbing call was produced with Avisoft-

569 SASLab.

570

571

572

573

574

575
576
577
578
579
580
581
582
583
584
585
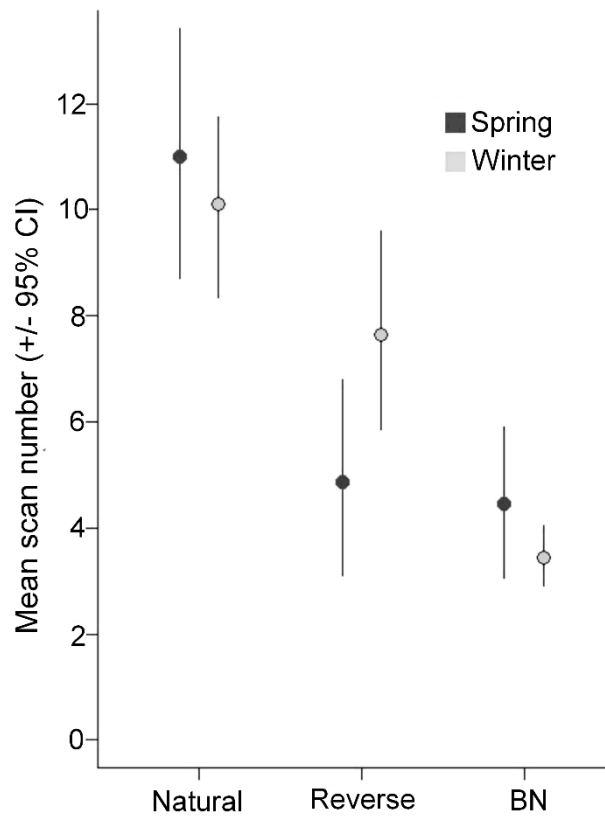
Figure 2. Mean number of scans (horizontal head movement of ~ 180°) produced in 1min by

Great Tits as a function of the call sequence used and the Season (± 95% boot-strapped

Confidence Intervals). Great Tit mobbing calls in their natural order (FME-D) are compared

to reversed calls (D-FME) and to Background Noise (BN). Statistical inference was made

using confidence intervals: difference is statistically significant when less than half of the CI

overlap (Cumming & Finch 2005, Cumming 2007).

597

598

599

600

601

602

603

604
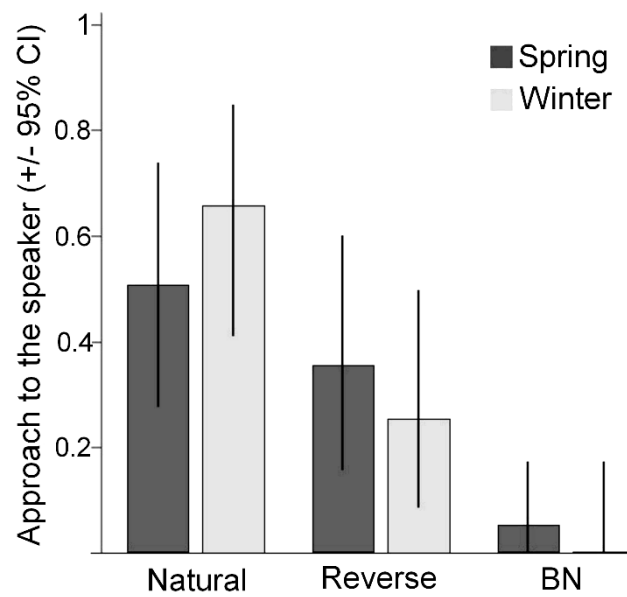
Figure 3. Percentage of individuals (± 95% CI) approaching the speaker as a function of the call sequence and the season. Great Tit mobbing calls in their natural order (FME-D) are compared to reversed calls (D-FME) and to Background Noise (BN). Statistical inference can be made using the confidence intervals: the difference is statistically significant when less than half of the CI overlap (Cumming & Finch 2005, Cumming 2007).