Functional flexibility in wild bonobo vocal behaviour

Zanna Clay, Jahmaira Archbold, Klaus Zuberbuhler

A shared principle in the evolution of language and the development of speech is the emergence of functional flexibility, the capacity of vocal signals to express a range of emotional states independently of context and biological function. Functional flexibility has recently been demonstrated in the vocalisations of pre-linguistic human infants, which has been contrasted to the functionally fixed vocal behaviour of non-human primates. Here, we revisited the presumed chasm in functional flexibility between human and non-human primate vocal behaviour, with a study on our closest living primate relatives, the bonobo (*Pan paniscus*). We found that wild bonobos use a specific call type ("peeps") across a range of contexts that cover the full valence range (positive-neutral-negative) in much of their daily activities, including feeding, travel, rest, aggression, alarm, nesting and grooming. Peeps were produced in functionally flexible ways in some contexts, but not others. Crucially, calls did not vary acoustically between neutral and positive contexts, suggesting that recipients take pragmatic information into account to make inferences about call meaning. In comparison, peeps during negative contexts were acoustically distinct. Our data suggest that the capacity for functional flexibility has evolutionary roots that predate the evolution of human speech. We interpret this evidence as an example of an evolutionary early transition away from fixed vocal signalling towards functional flexibility, representing the principle of human speech evolution.

1   Functional flexibility in wild bonobo vocal behaviour

2

3   Zanna Clay*[1,2], Jahmaira Archbold[1] & Klaus Zuberbühler[1, 3]

4

5   [1]* Department of Comparative Cognition, University of Neuchatel, Neuchatel, Switzerland; [2]

6   School of Psychology, University of Birmingham, UK; [3]School of Psychology &

7   Neurosciences, University of St Andrews, UK

8

9   *Z.Clay@bham.ac.uk; +44 7967 567 111; School of Psychology, University of Birmingham,

10   Birmingham, UK, B15 2TT

11

12   **Abstract**

13   A shared principle in the evolution of language and the development of speech is the

14   emergence of functional flexibility, the capacity of vocal signals to express a range of

15   emotional states independently of context and biological function. Functional flexibility has

16   recently been demonstrated in the vocalisations of pre-linguistic human infants, which has

17   been contrasted to the functionally fixed vocal behaviour of non-human primates. Here, we

18   revisited the presumed chasm in functional flexibility between human and non-human

19   primate vocal behaviour, with a study on our closest living primate relatives, the bonobo

20   (*Pan paniscus*). We found that wild bonobos use a specific call type ("peeps") across a range

21   of contexts that cover the full valence range (positive-neutral-negative) in much of their

22   daily activities, including feeding, travel, rest, aggression, alarm, nesting and grooming.

23   Peeps were produced in functionally flexible ways in some contexts, but not others.

24   Crucially, calls did not vary acoustically between neutral and positive contexts, suggesting

25   that recipients take pragmatic information into account to make inferences about call

26   meaning. In comparison, peeps during negative contexts were acoustically distinct. Our

27   data suggest that the capacity for functional flexibility has evolutionary roots that predate

28   the evolution of human speech. We interpret this evidence as an example of an

29   evolutionary early transition away from fixed vocal signalling towards functional flexibility,

30   representing the principle of human speech evolution.

31

32   **Introduction**

33   A growing body of research suggests that human infant vocal development reveals

34   something about the evolutionary history of language (Tomasello et al., 2005; Locke &

35   Bogin, 2006; Tomasello, 2008). This is because a basic principle of evolution is that natural

36   selection acts by modifying or adding complexity to existing structures and mechanisms

37   rather than by generating entirely new ones, a logic that has also been applied to the

38   evolution of language (Fitch, 2010). A research goal therefore is to describe the basic

39   design principles of early stages of vocal behaviour, which may have served as building

40   blocks on which subsequent stages of linguistic development have emerged (Oller, 2000;

41   Tomasello et al., 2005; Oller et al., 2013). One of these building blocks is 'functional

42   flexibility', an individual's capacity to produce signals that are detached from a

43   predetermined function to express different psychological states in a range of situations

44   (Griebel & Oller, 2008; Oller et al., 2013). This is thought to contrast with animal signals

45   and some human vocalisations (e.g. crying, laughter), which are tightly linked to specific

46   psychological and motivational states (Oller et al., 2013).

47

48   Part of this reasoning stems from a paucity of evidence of functionally flexible signalling in

49   non-human primates (Oller et al., 2013; Ackermann, Hage & Ziegler, 2014 but see Lameira

50   et al., 2013). In contrast, recent analyses of prelinguistic human infant vocal interactions

51   showed that three types of 'protophone' vocalisations were used flexibly to express a full

52   range of emotional content (positive, neutral and negative valence) across a range of

53   different situations (Oller et al., 2013). Moreover, some vocal types (squeals, vowel-like

54   sounds and growls) were associated with neutral facial expressions, further suggesting

55   that, in human infants, vocal production can be detached from a specific biological function.

56   This is in contrast to other infant vocalisations, namely crying and laughter, which appear

57   to be fixed to specific affective states. It has therefore been suggested that these types of

58   affect-bound, or functionally fixed, vocalisations resemble non-human primate calls, both in

59   terms of their form and function, as well as the brain mechanisms underlying them

60   (Newman, 1985; Owren, Amoss & Rendall, 2011; Bryant & Aktipis, 2014; Owren & Amoss,

61   2014).

62

63   In animals, functionally fixed signals are often considered equivalent to 'functionally

64   referential' signals, broadly defined as acoustically distinct signals produced in response to

65   a narrow range of stimuli to the extent that a receiver, upon hearing the signal in the

66   absence of the stimuli, responds *as if* experiencing the eliciting stimuli itself (Macedonia &

67   Evans, 1993). The alarm calls of numerous primates, for example, have been discussed as

68   being functionally fixed in that they are produced reliably in response to certain classes of

69   predators (such as aerial alarm calls in response to aerial predators) and their underlying

70   affective states (Seyfarth et al. 1980a, b; Zuberbühler, 2003, 2006). While some calls may fit

71   the notion of functional fixedness, the theoretical assumption that vocalisations must be

72   produced to a narrow range of stimuli in order to functionally refer to something in the

73   world has, in fact, been recently been challenged on the grounds that calls produced to a

74   number of different stimuli may still be functionally referential, given the availability of

75   other contextual cues (Scarantino & Clay, 2014).

76

77   More recent research on animal vocal signalling, particularly by non-human primates, is

78   increasingly suggesting scope for greater flexibility, the spontaneous production of

79   unvoiced, atypical calls by captive great apes being one such example. In many captive

80   facilities, there have been reports of individuals producing atypical, voiceless calls (e.g.

81   "raspberrys") in order to gain the attention of human caregiver, and that some of these are

82   acquired through social learning (e.g. *Gorilla gorilla* (Perlman, Patterson & Cohn, 2012);

83   *Pongo pygmaeus* (Wich et al., 2009; Hardus et al., 2009; Lameria et al., 2015); *Pan*

84   *troglodytes* (Hopkins, Taglialatela & Leavens, 2007; Taglialatela et al., 2012); *Pan paniscus*

85   (Taglialatela & Savage-Rumbaugh, 2003)). Moreover, conditioning experiments with

86   rhesus macaques and cotton-top tamarins have highlighted considerable vocal control,

87   with individuals able to initiate vocalisations and modify various features of their vocal

88   output in response to different external stimuli (Hage, Gavrilov & Nieder, 2013; Hotchkin,

89   Parks & Weiss, 2013). Studies of natural communication among conspecifics have also

90   highlighted considerable vocal flexibility. For instance, studies in the contexts of travel

91   recruitment (Gruber & Zuberbühler, 2013), feeding advertisement (Slocombe et al., 2010;

92   Schel et al., 2013), sexual interactions (Townsend, Deschner & Zuberbühler, 2008), social

93   greetings (Laporte & Zuberbühler, 2008) and predator alarms (Crockford et al., 2012; Schel

94   et al., 2013) have revealed that wild chimpanzees have a notable degree of control over call

95   production and flexibly modify it in response to different audience compositions and, in

96   some cases, in intentional ways (Crockford et al., 2012; Schel et al., 2013). Group-specific

97   'dialects' have also been suggested for four chimpanzee communities living in the same

98   forest in the Ivory Coast (Crockford et al., 2004). In orang-utans, a recent study of the

99   production of variants of an alarm call (the kiss-squeak) in two different wild populations

100   revealed population-specific usage, which indicated a certain degree of arbitrariness

101   (Lameira et al., 2013).

102

103   Research on the communicative behaviour of a human-enculturated and language-trained

104   bonobo ('Kanzi') has already highlighted bonobos as an interesting candidate species to

105   examine flexible primate vocal production (Savage-Rumbaugh & Lewin, 1994). It has been

106   suggested, for instance, that Kanzi used 'peeps', one of the most common vocalisations in

107   the bonobo repertoire (de Waal, 1988), to communicate with caregivers in ways that

108   conform to conversational rules used in human speech dialogue (Greenfield & Savage-

109   Rumbaugh, 1990, 1993; Savage-Rumbaugh, 1990, 1998; Segerdahl, Fields & Savage-

110   Rumbaugh, 2005). Another study also suggested that Kanzi's vocal repertoire had become

111   augmented to include novel acoustic variants of the 'peep' vocalisation produced in

112   response to different food types and that, astonishingly, the acoustic structure of these

113   vocal variants resembled the corresponding spoken English words used by his caregivers

114   (Hopkins & Savage-Rumbaugh, 1991; Taglialatela & Savage-Rumbaugh, 2003).

115

116   While Kanzi's communicative behaviour is suggestive of considerable communicative

117   flexibility, research into flexibility in the natural vocal communication system of bonobos

118   remains scant (Liebel et al., 2013). The bonobo vocal repertoire has been described as

119   highly graded, containing up to fifteen principal call types (de Waal, 1988; Bermejo &

120   Omedes, 1999). Consistent with the more traditional view of primate vocal signalling, most

121   bonobo call types appear to be tied to particular emotional states or valence classes, such

122   as pant-laughing during socio-positive interactions, screaming and pout-moaning in

123   response to agonism, threat barks during aggressive attacks, pant-grunts during

124   submissive greetings and alarm barks in response to predators (de Waal, 1988; Clay et al.,

125   Unpublished Data). However, as indicated in the Kanzi studies, their most frequent vocal

126   type, the 'peep', is produced across an array of behavioural contexts (Fig 1., (de Waal,

127   1988)). Analyses of the bonobo vocal repertoire in the wild (Bermejo & Omedes, 1999) and

128   in captivity (de Waal, 1988) both stressed the importance of the peep in bonobo

129   communication and reported its highly flexible and varied use across different behavioural

130   contexts.

131

132   Although peeps are produced across an array of contexts, they are especially common

133   during feeding events and, consequently, have so far only been systematically studied as a

134   food-associated signal (van Krunkelsven et al., 1996; Clay & Zuberbühler, 2009, 2011). For

135   instance, peeps produced in response to food discovery have been shown to be frequently

136   combined with other call types into longer sequences, whereby the probabilistic

137   organisation of the sequence conveys information to receivers about perceived food quality

138   (Clay & Zuberbühler, 2009, 2011).

139

140 The apparently varied usage of the peep vocalisation suggests it to be an interesting

141 candidate to study evidence for vocal flexibility but, to our knowledge, no systematic

142 investigation of their general use has been conducted. For example, it is currently unknown

143 whether peeps share the same acoustic structure across contexts or whether they are

144 context-specific. Before investigating the degree to which calls may be produced flexibly, it

145 is important to first establish whether its acoustic structure varies across contexts.

146

147 The aim of the current study is to systematically analyse the acoustic structure of peeps to

148 assess whether they are tied to specific behavioural contexts or whether, like human

149 infants (Oller et al., 2013), bonobos are capable of producing the same vocalisation across a

150 range of valence states. From an evolutionary perspective, looking for non-human primate

151 vocalisations that are not so tightly tied to biological function but are used in more

152 functionally flexible ways provides potentially relevant insights for the evolution of human

153 speech. If humans are the only species capable of functional flexibility, that is, to produce

154 the same vocalisation across different valence states, peeps should be expected to vary

155 according to the valence (i.e. from positive to neutral to negative) of the contexts. To this

156 end, we analysed various acoustic parameters of bonobo peeps produced in different

157 behavioural contexts relating to the three principal valence dimensions (positive-neutral-

158 negative) to explore whether these structural parameters provided acoustic cues relating

159 to the inferred affective valence. This approach differs from the 'discrete emotion'

160 approach, which presupposes discrete emotional states in animals and humans that arise

161 in response to anticipation of rewarding or punishing events, such as fear or pleasure

162 (Russell, 1980; Mendl, Burman & Paul, 2010). By taking a more graded approach, the

163 current study is better suited for making direct comparisons with previous research on

164 prelinguistic human infant vocalisations (Oller et al., 2013).

165

166 **Materials & Methods**

167

168 Behavioural observations were conducted on individuals from the Bompusa community of

169 wild bonobos by Z.C. from October 2013 to March 2014 at Lui Kotale, located near the

170 Salonga National Park, in DR Congo. At this time, the fully habituated and fully identified

171 community consisted of twelve adult females, two subadult females, five adult males two

172 subadult males and eighteen immatures (juveniles and infants). Animal focal animal

173 sampling (15 minutes) was conducted on all males and all adult females throughout the

174 study period, amounting to an average of 11 focal hours per individual.

175

176 We recorded vocalisations produced by focal individuals across a variety of behavioural

177 contexts, as they occurred. We recorded vocalizations at distances of 7–20 m using a

178 Sennheiser MKH816T directional microphone and Marantz PMD661 solid-state recorder

179 (Microphone frequency response: 50-20 000 Hz, ±3.5 dB; sampling rate of 44.1 kHz, 16 bits

180 accuracy).

181

182 Our acoustic analyses focussed on the bonobo peep, a high-frequency, closed mouth

183 vocalisation (approx. 2,200 Hz, de Waal, 1988, see Fig. 1), short in duration (approx. 0.1 sec

184 (de Waal, 1988, Clay & Zuberbühler, 2009)) and characterised by a simple, flat acoustic

185   form composed of several harmonics that are generally un-modulated. In order to analyse

186   the peep structure across different contexts, we first identified the most vocally active

187   individuals from focal recordings, identifying those that produced vocalisations in at least

188   two feeding contexts and two non-feeding contexts, which resulted in a sample of eight

189   individuals (four adult males, one subadult male and three adult females). Behavioural

190   contexts were mutually exclusive, i.e. peeps produced holding or consuming food while

191   travelling or resting were excluded. In order to compare the acoustic structure of peeps in

192   different contexts we compared the acoustic structure of peeps during the contexts that

193   generated the most peep vocalisations per individual. We collected peeps produced at the

194   onset of each behavioural context (i.e. at the onset of food discovery or travel). Because the

195   number of peeps produced at the vocal sequence onset varied across different calling

196   events, we analysed up the first three consecutive peeps produced at the beginning of a

197   vocal event by the same individual as this was the typically number of peeps produced in a

198   consecutive sequence. We calculated mean scores per parameter across the three peeps to

199   standardise across the same calling event.

200

201   In contrast to an analysis focussing on discrete emotional states, we were interested in first

202   establishing whether bonobo vocalisations may be used flexibly across the valence contexts

203   (positive-neutral-negative), as has been demonstrated in prelinguistic infants (Oller et al.,

204   2013). Therefore, for each individual, we randomly selected a balanced sample of eight

205   peep recordings produced during feeding contexts (feeding on shoots/seeds on the ground

206   and fruits in trees), which we inferred to be as approximately positive in overall valence

207   (Briefer, Tettamanti & McElligott, 2015) and eight peep events produced during non-

208   feeding contexts (resting and travel), which we inferred to be, in comparison to feeding,

209   relatively neutral in valence. In order to capture the spectrum of emotional valence in our

210   acoustic analyses (i.e. positive-negative-neutral), we also analysed a sample of peeps

211   associated with predator alarm responses and in response to agonistic interactions as the

212   victim, which were both taken to represent negative valence. As peeps in response to

213   agonistic and alarm contexts were rare, we analysed a balanced and randomized sample of

214   4 peep samples per individual, taken from independent behavioural events produced by 7

215   of the original 8 individuals (N = 28 in total). The eighth individual was excluded in this

216   sample due to inadequate sample size.

217   This amounted to a total of 128 peep events. We selected two contexts per valence class

218   (positive-negative-neutral) in order to maximise sample size as well as to adequately

219   capture the potential acoustic variation in different contexts. In order to capture variation

220   in the feeding experience overall, recordings from feeding contexts included a randomized

221   and balanced selection of vocal events in response to feeding on fruits in trees as well as to

222   herbaceous shoots on the ground. For non-feeding contexts, we analysed a randomized

223   sample of recordings for each individual produced during rest and travel on the ground.

224   For negative valence contexts, we analysed a randomised balanced sample of peeps

225   produced during agonistic conflicts and predator alarm contexts.

226

227   We carried out all quantitative acoustic analyses using Praat 5.4.01 using the following

228   settings: analysis window length 0.05 s, dynamic range 70 dB; pitch range 500–3000 Hz,

229   optimized for voice analysis, spectrogram view range 0–10 kHz. We performed pitch

230   analysis using a script ("Analyse Source Editor") written by M. Owren (personal

231   communication). We then took the following spectral measurements from the fundamental

232   frequency (F0): 1) mean fundamental frequency (Hz): average F0 across the entire call; (2)

233   frequency at call onset, (3) frequency at call middle; (4) frequency at call offset; (5)

234   transition onset (Hz): frequency of maximum energy at call onset minus frequency of

235   maximum energy at call middle; (6) transition offset (Hz): frequency of maximum energy at

236   call middle minus frequency of maximum energy at call offset; (7) maximum fundamental

237   frequency (Hz): maximum frequency of F0; (8) minimum fundamental frequency (Hz):

238   minimum frequency of F0; (9) number of harmonics: number of harmonic bands visible. In

239   the temporal domain, we measured the call duration (10).

240

241   Next, we screened the data for outliers by producing standardized Z scores, rejecting any

242   calls with a Z score greater than 3.29 in one or more parameters (Tabachnick & Fidell,

243   2001). We regressed all parameters to check for multi-colinearity and singularity,

244   removing parameters with a variance inflation factor greater than 10. We then conducted a

245   discriminant function analysis (DFA) to assess whether the uncorrelated acoustic variables

246   could discriminate between different behavioural contexts. Each of the eight individuals

247   equally contributed eight randomly selected calls for both food (henceforth 'positive

248   valence') and non-food (henceforth 'neutral valence') contexts and four calls per individual

249   were entered for the negative valence (N = 156 peep samples in total). To cross-validate

250   the discriminant functions produced in the analysis, we used the leave-one-out

251   classification procedure, which classifies each calls by the functions derived from all calls

252   other than that one. We used Binomial tests to analyse whether the proportion of correct

253   discrimination differed significantly from chance.

254

255   In order to examine whether peeps conveyed information about caller identity, we

256   conducted a DFA using the same data used for the above analysis but taking individual

257   identity as the discriminating factor. We additionally conducted separate DFAs for the

258   positive and neutral valence contexts in order to control for behavioural context. We were

259   unable to include separate DFAs for the negative valence context due to small sample size

260   (N = 4 calls per individual) compared to the number of acoustic parameters under scrutiny,

261   which led to inadequate statistical power.

262

263   Since the acoustic data were two-factorial (caller ID; context), it has been argued that

264   conventional DFA does not allow for a valid estimation of the overall significance of

265   discriminability (Mundry & Sommer, 2007). Therefore, for any significant DFA

266   discrimination, we conducted a permuted Discriminant Function Analysis (pDFA), using a

267   macro written by R. Mundry and C. Sommer (Mundry & Sommer, 2007; R. Mundry,

268   Personal Communication). The pDFA estimates the significance of the number of correctly

269   classified calls (cross-validated), taking into account repeated contributions per individual

270   caller.

271

272   Following significant discrimination in the pDFA and diagnostic tests, we used Univariate

273   Analyses of Variance tests to explore whether each of the acoustic parameters varied

274   statistically with context, entering Caller Identity as a Random Factor and Context as the

275   Fixed Factor.

276

277 All statistical tests were carried out using SPSS version 21.0 (SPSS Inc., Cary, NC, U.S.A.) and

278 R Studio version 3.1.1 (The R Foundation for Statistical Computing, Vienna, Austria). All

279 tests were two tailed and alpha levels were set at 0.05, unless stated as being corrected. We

280 applied standarised Bonferroni corrections for multiple comparisons.

281

282 **Results**

283 During focal animal sampling, we recorded peeps in response to over a dozen different

284 behavioural contexts, which, across all focal individuals, included feeding on fruits, leaves,

285 seeds, flowers in trees and on shoots, seeds, leaves and fruits on the ground. It also

286 included travelling, resting, grooming, preparing a nest, interacting sexually, responding to

287 vocalisations from other parties, descending from trees after feeding, alarm responses to

288 predators or unexpected events, weather changes, agonistic interactions, submissive or

289 appeasement responses towards more dominant individuals, and vocal greetings to the

290 arrival of another individual into the party.

291

292 *Acoustic structure of peeps*

293 We compared the acoustic structure of peeps produced in different contexts (feeding;

294 travel/rest; agonism/alarm) that were associated with different emotion valences

295 (positive; neutral; negative valence, respectively), Fig. 1 and that generated the most peep

296 vocalisations across individuals. Following a multi-colinearity screening, we entered six of

297 the nine original acoustic parameters into our acoustic analyses for eight individuals (total

298 N call events = 156: call duration, mean F0, F0 at call onset, number of harmonics,

299 transition onset and transition offset) and applied logarithmic transformations on three of

300   the acoustic parameters to improve their homogeneity of variance. Results from a cross-

301   validated discriminant function analysis revealed that while the DFA model generated two

302   significant discriminant functions (Wilks Lambda: 0.550, χ^2 (df = 14) = 80.007, P < .001),

303   peeps produced in association with positive valence contexts could not be reliably

304   discriminated from those produced in all other contexts: the functions only classified

305   49.3% of the calls correctly, which was below chance level (Binomial test (0.14) P > 0.05).

306

307   On a pairwise basis, DFA analyses further revealed that peeps produced in association with

308   positive valence contexts could not be reliably discriminated from those produced during

309   neutral valence contexts (Wilks Lambda: 0.947, χ^2 (df = 6) = 6.638, P = 0.356). In a cross-

310   validated analysis, the functions only classified 52.3% of the calls correctly, which was

311   below chance level (Binomial test (0.5) P > 0.05). However, peeps associated with negative

312   valence (i.e. alarm and agonism) could be significantly discriminated from those produced

313   in association with feeding (82.1% of calls correctly classified; Wilks lambda = 0.468, χ^2 (df

314   = 7) = 59.602, P < .001; Binomial test (0.5) P < .001, Bonferroni corrections), which was

315   validated in a subsequent pDFA controlling for repeated contributions (P = 0.009).

316   Similarly, there was significant discrimination of peeps produced in response to negative

317   valence contexts to those during neutral valence contexts, with 77.4% of calls (cross-

318   validated) correctly classified (Wilks Lambda = 0.551, χ^2 (df = 6) = 47.107, P < .001;

319   Binomial test (0.5) P < .001, Bonferroni corrections).

320

321   *Caller Identity*

322   We used the same cross-validated DFA procedure to test whether peeps could be

323   acoustically discriminated on the basis of caller identity (N = 8 individuals). The model

324   generated six significant discriminant functions (Wilks Lambda: 0.371, χ^2 (df = 42) =

325   119.043, P < .001), which discriminated caller identity at a significantly higher rate than

326   chance (cross-validated correct classification: 31.3%, Binomial test (0.125) P < 0.001).

327

328   We then conducted two separate DFAs to examine individual identity discrimination for

329   peeps in positive and neutral contexts. Results from the analyses were equivalent, with

330   identity significantly discriminated in both contexts (Individual identity in Feeding

331   contexts 31.3% (20/64) calls correctly classified: Wilks Lambda = 0.234, χ^2 (df = 42) =

332   81.285, P < 0.001; Binomial test (0.125) p <0.001; in non-feeding contexts: 32.8% (21/64)

333   calls correctly classified Wilks lambda = 0.210, χ^2 (df = 42) = 87.313, P < 0.001; Binomial

334   test (0.125) P < .001).

335

336   *Comparing acoustic parameters*

337   At the level of acoustic parameters, Univariate ANOVAs (Caller Identity as a random factor)

338   revealed that the mean call duration, the mean fundamental frequency and the mean

339   frequency at call onset varied significantly as a function of behavioural context (Mean call

340   duration $F_{2,12}$ = 5.625, P = 0.019; Mean F0: $F_{2,12}$ = 19.054, P < .001; F0 call onset: $F_{2,12}$ =

341   40.259, P < 0.001). Pair-wise comparisons (standard Bonferroni corrections), as shown in

342   Figure 2, of fundamental frequency (F0) parameters showed that peeps produced in

343   association with negative valence had a significantly higher mean F0 and a higher onset F0

344   compared to peeps associated with positive valence (Mean F0$_{negative}$ = 2131 ± 267 Hz, Mean

345   F0$_{positive}$ = 1660 ± 133; $F_{1,6}$ = 16.862, P = 0.006; F0 at call onset$_{negative}$ = 2027 ± 194 Hz, F0 at

346   call onset$_{positive}$ = 1612 ± 125 Hz; $F_{1,6}$, = 35.990, P = 0.001) and neutral valence (Mean

347   F0$_{neutral}$ = 1584 ± 210 Hz: $F_{1,6}$ = 27.160, p = 0.002; F0 at call onset$_{neutral}$l 1508 ± 186 Hz, $F_{1,6}$ =

348   69.887, P < .001). Although peeps associated with negative valence were shorter in

349   duration compared to those associated with positive valence (Mean call duration$_{negative}$ =

350   0.12 ± 0.14, mean call duration$_{positive}$ = 0.15 ± 0.03: $F_{1,6}$ = 8.316, P = 0.028), the result was

351   not significant under the Bonferroni correction. There were no other significant acoustic

352   differences.

353

354   **Discussion**

355   Results suggest that, contrary to current models, humans may not be unique among

356   primates in their ability to produce functionally flexible vocalisations. Our acoustic

357   analyses suggest that the 'peep' calls of wild bonobos are produced in such flexible ways in

358   response to a range of different behavioural contexts of varying affective valence.

359   Following similar evidence in pre-linguistic human infants (Oller et al., 2013), bonobos

360   produced peep vocalisations across a range of neutral, negative and positive contexts in

361   different behavioural situations. Although peeps produced in negative contexts differed

362   acoustically, the acoustic structure of peeps produced in association with positive valence

363   (feeding), could not be discriminated from neutral valence (travel, rest), despite the fact

364   that behavioural contexts were mutually exclusive. Specifically, peeps produced in

365   association with negative valence (alarm and agonism) possessed significantly higher mean

366   fundamental frequencies, higher frequencies at call onset and shorter durations. The

367   finding of acoustic variants within the same call type across different contexts is a relatively

368   common finding in studies of primate vocal behaviour (Owren et al., 1997; Rendall et al.,

369   1999; Rendall, 2003; Slocombe & Zuberbuhler, 2007). Moreover, the fact that context-

370   specific acoustic cues were found in negative but not other contexts has interesting

371   implications both at the level of call production and evolutionary function. In terms of call

372   production, the acoustic differences (higher frequencies and shorter call duration) found

373   for negative contexts are most likely a direct result of higher subglottal air pressure during

374   call production (Fitch & Hauser, 1995; Fitch, 2006), probably because individuals perceive

375   situations as charged, tense and urgent. This is likely to have direct physiological

376   consequences such that the same volume of air is passing through the vocal tract but at a

377   higher speed so that the vocal folds oscillate at a higher frequency for a shorter period of

378   time.

379

380   The fact that peeps produced during negative contexts are more constrained by the

381   physical mechanics of vocal production raises the possibility that, in the course of language

382   evolution, functional flexibility may first have occurred in positive and neutral contexts.

383   Generally, cues to emotional valence are mechanistically harder to conceal during negative

384   contexts, which impedes the vocal control that is required for functional flexibility.

385   Currently, it is not known whether the functionally flexible calls of human infants also vary

386   acoustically as a function of affect valence, as no acoustic analyses have been conducted

387   (Oller et al., 2013). Our prediction is that, even in humans, negative emotional valence is

388   equally conveyed in the acoustic structure of human vocalisations, including infant calls.

389

390   Overall, our results suggest that bonobo peep production may represent a somewhat

391   intermediate stage between functionally fixed (as seen for most primate vocalisations) and

392   functionally flexible signals, as seen in most human vocalisations (Oller et al., 2013).

393   Nevertheless, peeps could be reliably discriminated on the basis of caller identity across

394   contexts, suggesting that the calls reliably convey other relevant information to receivers.

395

396   In order to make comparisons with evidence from pre-linguistic human infants (Oller et al.,

397   2013), our study took a 'dimensional approach' to emotions (Russell, 1980; Watson et al.,

398   1999; Mendl, Burman & Paul, 2010) by focussing on the valence associated with the

399   eliciting behavioural contexts. According to the 'dimensional approach', emotions are

400   characterized along two core dimensions: valence (negative or positive; e.g. sad versus

401   happy; Russell, 1980) and arousal (e.g. calm versus excited (Mendl, Burman & Paul, 2010).

402   This contrasts with the 'discrete emotions' approach (Ekman, 1992; Panksepp, 1998),

403   which although valuable for examining specific emotions, has been suggested to over-focus

404   on certain emotions and not others (e.g. positive ones (Boissy et al., 2007), as well as

405   lacking an overall framework that integrates a wider range of emotional states (Mendl,

406   Burman & Paul, 2010). The 'dimensional approach' has been shown to be more powerful

407   for studying animal emotions (Mendl et al., 2009; Mendl, Burman & Paul, 2010) across

408   behavioural (Briefer, Tettamanti & McElligott, 2015; Reefman, Weschler & Gygax, 2009;

409   Imfeld-Mueller et al., 2011), physiological (da Costa et al., 2004; Davies, Radford & Nicol,

410   2014), and cognitive domains (Nettle & Bateson, 2012; Briefer & McElligott, 2013).

411   Consequently, the current study is therefore unable to provide more detailed insights into

412   the relationship between vocal production and discrete emotional states, such as fear or

413   pleasure (Ekman, 1992; Panksepp, 1998). It can be assumed that considerable variation in

414   discrete emotional states most likely exists within the contexts recorded here, for instance,

415   feeding may elicit different discrete emotional states according to other related factors (e.g.

416   social rank, age, group size). A substantial follow-up study, involving a greater sample size

417   of both calls and contexts, that also controls for potentially confounding factors, would be

418   needed to address the vocal indicators of discrete emotional states.

419

420   It has been suggested that the presence of functionally flexible vocalisations in pre-

421   linguistic human infants is evidence for an evolutionary divergence towards speech

422   production and variation in the expression of emotion across different utterances that sets

423   humans apart from the rest of the primate lineage (Oller et al., 2013). For instance, these

424   vocalisations, known as 'protophones' are produced across a full range of affect states and

425   were the most commonly occurring vocal type, suggesting that even pre-linguistic infants

426   possess considerable 'exploratory vocal freedom' and utilise a communication system that

427   is predominantly detached from function. Protophones were most commonly associated

428   with neutral facial expressions compared to cry and laughter, which were tightly linked to

429   negative and positive affect, respectively, further highlighting their greater emotional

430   detachment. Interestingly, bonobo peeps are also typically produced with a neutral facial

431   expression, produced while the mouth remains closed without any particular facial

432   expression (de Waal, 1987).

433

434   As always, it is somewhat problematic to draw firm conclusions from negative evidence,

435   that is, the fact that we found no acoustic differences across non-negative valence contexts.

436   For example, it is possible that there was subtle variation in other acoustic parameters that

437   we did not analyse, such as amplitude modulation (Fichtel & Hammerschmidt, 2001). Our

438   recordings were made under natural conditions with free-ranging animals, which led to

439   various constraints, such as recording at varying distances, directions and atmospheric

440   conditions, which made it impossible to derive reliable measures of amplitude modulation.

441   Future research on peeps using a larger sample size in more controlled acoustic conditions

442   could address this issue. More conclusive evidence would require carefully designed

443   playback experiments to test whether bonobos have more difficulties discriminating

444   positive and neutral peeps, compared to negative peeps. This is important because the

445   notion of functional flexibility also makes predictions about signaller-receiver interactions,

446   i.e. between a signaller's communicative act, or 'illocutionary force', and a recipient's

447   consequential response, or 'perlocutionary effect'. For example, an infant's "complaints"

448   may be associated with specific vocalisations showing negative affect, which then induces a

449   caregiver offering to feed after interpreting the negative affect vocalisation as an indicator

450   of hunger (Austin, 1975; Bates et al., 1979, Oller et al. 2013). The interaction between

451   signaller and receiver would be a relevant aspect to explore in future work, such as using

452   match to sample experiments or naturalistic playback studies.

453

454   While functional flexibility is undoubtedly more developed in human speech, results from

455   the current study contribute to the debate surrounding speech evolution by highlighting

456   evidence for a possible intermediate stage in the communicative repertoire of our closest

457   living relative, the bonobo. While much of bonobo vocal behaviour is functionally fixed by

458   nature, such as screaming during agonistic encounters, barking during alarm and laughing

459   during play (de Waal, 1988), their most common vocalisation, the 'peep' is also produced in

460   highly flexible ways across a wide range of contexts and valence states. Previously, the

461   peep vocalisation had been primarily considered as a food-associated call that was

462   combined with other calls in a sequence in order to convey information about food quality

463   (de Waal, 1988; Clay & Zuberbühler, 2009, 2011). It is possible that when combined in a

464   particular way with other calls, or when combined with a certain context, peeps may still

465   functionally refer to specific events in the external world (Scarantino & Clay, 2014). Future

466   research, probing the manner in which peeps are combined with other calls and contextual

467   stimuli, is needed to address this question.

468

469   It is relevant to note that studies of language-competent bonobos have already highlighted

470   bonobos as a good model for studying the prerequisites for language evolution (Hopkins &

471   Savage-Rumbaugh, 1991; Taglialatela & Savage-Rumbaugh, 2003). In addition to their

472   remarkable ability to acquire human and artificial language systems in captivity (e.g.

473   Segerdahl, Fields, & Savage-Rumbaugh, 2005), bonobo vocal behaviour has been suggested

474   to be more flexible and dialogue-like than chimpanzees. An interesting possibility is that

475   peeps function to draw attention to and "comment" on novel items or environmental

476   events (de Waal, 1988; Savage-Rumbaugh, Shanker & Taylor, 1998), a communicative

477   behaviour also found in early language development (Tomasello & Carpenter, 2007).

478   Future research should determine whether bonobos are also capable of varying the

479   expression of emotional valence across different utterances within the same vocal event

480   using the same vocal type.

481

482   An alternative possibility is that the production of peeps fits within a more generalised

483   function that extends across multiple contexts (Notman & Rendall, 2005). For instance, it

484   may have a broader social function related to cohesion or spacing which may explain

485   patterns of usage across seemingly disparate behavioural contexts. For example, the

486   vocalisations produced during feeding by dolphins (Janik, 2003) and greater-spear nosed

487   bats (Wilkinson & Boughman, 1998) have been suggested to function to coordinate social

488   foraging rather than being specific to feeding per se. The rationale for cohesion effects may

489   even be supported by the current results regarding negative contexts, since cohesion in this

490   case could be outweighed by the urge to escape a predator or a conspecific.

491

492   Another proximate explanation for shared acoustic structure is an underlying a shared

493   motivational aspect to the eliciting contexts in question. To address this issue, a solid

494   framework to assess expressions of affect in great apes is first required. One approach,

495   already employed for human infants (Oller et al., 2013), could involve comparing affect

496   expression in corresponding modalities, such as coding facial expressions produced in

497   association with the vocalisation (e.g. Parr et al., 2008; Parr, Waller & Vicks, 2007).

498   However, unlike the facial affect analyses conducted in association with vocalisation

499   production for human infants (Oller et al., 2013), bonobo peep vocalisations are closed-

500   mouth vocalisations, making facial affect coding less appropriate. Future work using

501   associated playbacks and hormonal analyses may also provide relevant insights into

502   underlying motivational states during vocal production (Mateo, 2010).

503

504   In summary, the current study contributes promising insights into the evolution of human

505   speech by suggesting an intermediate stage between fully-fledged functional flexibility in

506   human speech production and the more traditionally viewed of fixed signals of non-human

507   primates. Rather than being a uniquely human trait, results from the current study reflect a

508   growing body of literature that suggests that flexible vocal signalling (e.g. Lemasson et al.,

509   2005, 2011; Ouattara, Lemasson & Zuberbühler, 2009; Crockford et al., 2012; Koda et al.,

510   2013; Liebel et al., 2013) and perhaps even functional arbitrariness (Lameria et al., 2013)

511   may have deeper roots in the primate lineage than previously assumed. By demonstrating

512   the potential for functional flexibility in the vocal behaviour of a great ape species, the

513   results provide a useful springboard from which evidence for functional flexibility in other

514   animal species and across different modalities can be explored.

515

530

References

532

Ackermann H, Hage SR, & Ziegler W. (2014). Brain mechanisms of acoustic communication in humans and nonhuman primates: An evolutionary perspective. *Behavioral and Brain Sciences, 37*(06): 529-546.

Austin JL. (1975). *How to do things with words* (Vol. 367). Oxford: Oxford University Press.

Bates E, Benigni L, Bretherton I, Camaioni L & Voltera V (1979). *The emergence of symbols: Cognition and communication in infancy*. New York: Academic Press.

Bermejo M, & Omedes A. (1999). Preliminary vocal repertoire and vocal communication of wild bonobos (*Pan paniscus*) at Lilungu (Democratic Republic of Congo). *Folia Primatologica, 70*(6): 328-357.

Boissy A, Manteuffel G, Jensen MB, Oppermann MR, Spruijit B, Keeling LJ, Winckler C, Forkman B, Dimitrov I, Langbein J, Bakken M, Veissier I & Aubert A. (2007) Assessment of positive emotions in animals to improve their welfare. *Physiology and Behaviour* 92(3): 375-397.

Briefer EF & McElligott AG. (2013). Rescued goats at a sanctuary display positive mood after former neglect. *Applied Animal Behavioral Science* 146(1):45-55

Briefer EF, Tettamanti F & McElligott AG. (2015). Emotions in goats: mapping physiological, behavioural and vocal profiles. *Animal Behaviour, 99*: 131-143.

Bryant GA, & Aktipis CA. (2014). The animal nature of spontaneous human laughter. *Evolution and Human Behavior, 35*(4): 327-335.

552   Clay Z, Ravaux L, de Waal FBM, Zuberbuhler, K . *Under Review*. Bonobos vocally protest to

553   violations of social expectations.

554   Clay Z, & Zuberbühler K. (2009). Food-associated calling sequences in bonobos. *Animal*

555   *Behaviour, 77*(6): 1387-1396.

556   Clay Z, & Zuberbühler K. (2011). Bonobos extract meaning from call sequences. *PLoS One,*

557   *6*(4): e18786.

558   Crockford C, Herbinger I, Vigilant L, & Boesch C (2004) Wild chimpanzees produce group-

559   specific calls: a case for vocal learning? *Ethology* 110(3): 221-243.

560   Crockford C, Wittig RM, Mundry R, & Zuberbühler K. (2012). Wild chimpanzees inform

561   ignorant group members of danger. *Current Biology, 22*(2): 142-146.

562   Da Costa AP, Leigh AE, Man M-S, & Kendrick KM (2004) Face pictures reduce behavioural,

563   autonomic, endocrine and neural indices of stress and fear in sheep. *Proceedings of*

564   *the Royal Society of London B: Biological Sciences* 271(1552): 2077-2084.

565   Davies A, Radford A, & Nicol C (2014) Behavioural and physiological expression of arousal

566   during decision-making in laying hens. *Physiology and Behaviour* 123:93-99.

567   de Waal FBM. (1988). The communicative repertoire of captive bonobos (*Pan paniscus)*

568   compared to that of chimpanzees. *Behaviour, 106*(3-4): 183-251.

569   Ekman P (1992) Are there basic emotions? *Psycholgical Review*, 99: 550-553

570   Fichtel C, & Hammerschmidt K. (2001). On the vocal expression of emotion. A multi-

571   parametric analysis of different states of aversion in the squirrel monkey. *Behaviour*

572   *138* (1): 97-116.

573   Fitch WT. (2006). Production of vocalizations in mammals. *Visual Communication, 3*: 145.

574   Fitch WT, & Hauser MD. (1995). Vocal Production in Nonhuman-Primates - Acoustics,

575   physiology, and functional constraints on honest advertisement." *American Journal*

576   *of Primatology, 37*(3): 191-219.

577   Greenfield PM, & Savage-Rumbaugh ES. (1990). Grammatical combination in *Pan paniscus*:

578   Processes of learning and invention in the evolution and development of language.

579   In S. T. Parker & K. R. Gibson, eds. *'Language' and intelligence in monkeys and apes*.

580   Cambridge: Cambridge University Press, 540-578.

581   Greenfield PM. & Savage-Rumbaugh, ES. (1993). Comparing communicative competence in

582   child and chimp: the pragmatics of repetition. *Journal of Child Language, 20*(01): 1-

583   26.

584   Griebel U & Oller D (2008) Evolutionary forces favoring communicative flexibility. In: Oller

585   DK, Griebel U, eds. *Evolution of communicative flexibility: Complexity, creativity, and*

586   *adaptability in human and animal communication*. Cambridge, MA: MIT Press, 9-40.

587   Gruber T & Zuberbühler K (2013) Vocal Recruitment for Joint Travel in Wild Chimpanzees.

588   *PloS One* 8(9):e76073.

589   Hage SR, Gavrilov N, & Nieder A. (2013). Cognitive control of distinct vocalizations in

590   rhesus monkeys. *Journal of Cognitive Neuroscience, 25*(10): 1692-1701.

591   Hardus ME, Lameira AR, Van Schaik CP, & Wich SA. (2009). Tool use in wild orang-utans

592   modifies sound production: a functionally deceptive innovation? *Proceedings of the*

593   *Royal Society B: Biological Sciences*, 1027.

594   Hopkins WD, & Savage-Rumbaugh ES. (1991). Vocal communication as a function of

595   differential rearing experiences in *Pan paniscus*: A preliminary report. *International*

596   *Journal of Primatology, 12*(6): 559-583.

597   Hopkins WD, Taglialatela JP, & Leavens DA. (2007). Chimpanzees differentially produce

598   novel vocalizations to capture the attention of a human. *Animal Behaviour, 73*(2):

599   281-286.

600   Hotchkin CF, Parks SE, & Weiss DJ. (2013). *Vocal modifications in primates: Effects of noise*

601   *and behavioral context on vocalization structure.* Proceedings of Meetings on

602   Acoustics (Acoustical Society of America), p 010061.

603   Imfeld-Mueller S, Van Wezemael L, Stauffacher M, Gygax L, & Hillmann E (2011) Do pigs

604   distinguish between situations of different emotional valences during anticipation?

605   *Applied Animal Behavioural Science* 131(3):86-93.

606   Janik V. (2000). Food–related bray calls in wild bottlenose dolphins (*Tursiops truncatus*).

607   *Proceedings of the Royal Society of London. Series B: Biological Sciences, 267*(1446):

608   923-927.

609   Koda H, Lemasson A, Oyakawa C, Pamungkas J, & Masataka N. (2013). Possible role of

610   mother-daughter vocal interactions on the development of species-specific song in

611   gibbons. *PloS One, 8*(8): e71432.

612   Lameira AR, Hardus ME, Bartlett AM, Shumaker RW, Wich SA, & Menken SB. (2015).

613   Speech-Like Rhythm in a Voiced and Voiceless Orangutan Call. *PloS One, 10*(1):

614   e116136.

615   Lameira AR, Hardus ME, Nouwen KJ, Topelberg E, Delgado RA, Spruijt BM, Wich SA. (2013).

616   Population-Specific Use of the Same Tool-Assisted Alarm Call between Two Wild

617   Orangutan Populations (*Pongopygmaeus wurmbii*) Indicates Functional

618   Arbitrariness. *PloS One, 8*(7): e69749.

619   Laporte MN & Zuberbühler K (2010) Vocal greeting behaviour in wild chimpanzee females.

620   *Animal Behaviour* 80(3):467-473.

621   Lemasson A, Hausberger M, & Zuberbühler K. (2005). Socially meaningful vocal plasticity in

622   adult Campbell's monkeys (*Cercopithecus campbelli*). *Journal of Comparative*

623   *Psychology, 119*(2): 220.

624   Lemasson A, Ouattara K, Petit EJ, & Zuberbühler K. (2011). Social learning of vocal

625   structure in a nonhuman primate? *BMC Evolutionary Biology, 11*(1): 362.

626   Liebal K, Waller BM, Slocombe KE, & Burrows AM (2013) *Primate Communication: a*

627   *multimodal approach*. Cambridge: Cambridge University Press.

628   Locke JL, & Bogin B. (2006). Language and life history: A new perspective on the

629   development and evolution of human language. *Behavioral and Brain Sciences,*

630   *29*(3): 259-279.

631   Macedonia JM, & Evans CS. (1993). Variation among mammalian alarm call systems and the

632   problem of meaning in animal signals. *Ethology, 93*: 177-197.

633   Mateo JM (2010) Alarm calls elicit predator-specific physiological responses. *Biology*

634   *Letters* :rsbl20100118.

635   Mendl M, Burman OH, Parker RM, & Paul ES (2009) Cognitive bias as an indicator of animal

636   emotion and welfare: emerging evidence and underlying mechanisms. *Applied*

637   *Animal Behavioral Science* 118(3):161-181.

638   Mendl M, Burman OH, & Paul ES (2010) An integrative and functional framework for the

639   study of animal emotion and mood. *Proceedings of the Royal Society B: Biological*

640   *Sciences* 277(1696):2895-2904.

641   Mundry R, & Sommer C. (2007). Discriminant function analysis with nonindependent data:

642   consequences and an alternative. *Animal Behaviour, 74*(4): 965-976.

643   Nettle D & Bateson M. (2012) The evolutionary origins of mood and its disorders. *Current*

644   *Biology* 22(17):R712-R721.

645   Newman JD. (1985). The infant cry of primates. In: Lester BM, Zachariah Boukydis CF., eds.,

646   *Infant crying*. Springer:New York, 307-323.

647   Notman H, & Rendall D. (2005). Contextual variation in chimpanzee pant hoots and its

648   implications for referential communication. *Animal Behaviour, 70*, 177-190.

649   Oller DK (2000). *The emergence of the speech capacity*. Mahwah, NJ: Lawrence Erlbaum

650   Associates.

651   Oller DK, Buder EH, Ramsdell HL, Warlaumont AS, Chorna L, & Bakeman R. (2013).

652   Functional flexibility of infant vocalization and the emergence of language.

653   *Proceedings of the National Academy of Sciences, 110*(16), 6318-6323.

654   Ouattara K, Lemasson A, & Zuberbühler K. (2009). Campbell's monkeys concatenate

655   vocalizations into context-specific call sequences. *Proceedings of the National*

656   *Academy of Sciences, 106*(51), 22026.

657   Owren MJ, & Amoss RT. (2014). Spontaneous human laughter. In: Tugade MM, Shiota, MN,

658   Kirby LD, eds. *Handbook of Positive Emotions*. New York: Guildford Press, 159-178.

659   Owren MJ, Amoss RT, & Rendall D. (2011). Two organizing principles of vocal production:

660   Implications for nonhuman and human primates. *American Journal of Primatology,*

661   *73*(6), 530-544.

662   Owren MJ, Seyfarth RM & Cheney DL. (1997). The acoustic features of vowel-like grunt

663   calls in chacma baboons (*Papio cynocephalus ursinus*): implications for production

664   processes and functions. *Journal of the Acoustical Society of America*, 101: 2951–

665   2963.

666   Panksepp J (1998) *Affective neuroscience: The foundations of human and animal emotions*.

667   New York, NY: Oxford University Press.

668   Parr LA, Waller BM, & Heintz M. (2008). Facial expression categorization by chimpanzees

669   using standardized stimuli. *Emotion, 8*(2): 216.

670   Parr LA, Waller BM, & Vick SJ. (2007). New developments in understanding emotional facial

671   signals in chimpanzees. *Current Directions in Psychological Science, 16*(3), 117-122.

672   Perlman M, Patterson FG, & Cohn RH. (2012). The human-fostered gorilla Koko shows

673   breath control in play with wind instruments. *Biolinguistics, 6*(3-4): 433-444.

674   Reefmann N, Wechsler B, & Gygax L (2009) Behavioural and physiological assessment of

675   positive and negative emotion in sheep. *Animal Behaviour* 78(3):651-659.

676   Rendall D. (2003) Acoustic correlates of caller identity and affect intensity in the vowel-like

677   grunt vocalizations of baboons. *Journal of the Acoustical Society of America*, 113 (6):

678   3390-3402

679   Rendall, D, Seyfarth, RM, Cheney DL & Owren, MJ (1999) The meaning and function of

680   grunt variants in baboons. *Animal Behaviour* 57: 583-592.

681   Russell JA. (1980). A circumplex model of affect. *Journal of Personality and Social*

682   *Psychology,* 39(6):1161

683   Russell JA & Barrett LF (1999) Core affect, prototypical emotional episodes, and other

684   things called emotion: dissecting the elephant. *Journal of Personality and Social*

685   *Psychology*, 76(5):805.

686   Savage-Rumbaugh ES. (1988). A new look at ape language: Comprehension of vocal speech

687   and syntax. Comparative perspectives in modern psychology. In D. W. Leger, ed.

688   *Nebraska Symposium on Motivation*. Lincoln, NB: University of Nebraska Press. Vol.

689   35, 201-256.

690   Savage-Rumbaugh ES. (1990). Language acquisition in a nonhuman species: Implications

691   for the innateness debate. Special Issue: The idea of innateness: Effects on language

692   and communication research. *Developmental Psychobiology, 23*(7), 599-620.

693   Savage-Rumbaugh ES, Shanker SG, & Taylor TJ. (1998). *Apes, language, and the human*

694   *mind*. New York: Oxford University Press.

695   Savage-Rumbaugh ES, & Lewin R. (1994). *Kanzi. The ape at the brink of the human mind*.

696   New York: Wiley.

697   Scarantino A, & Clay Z. (2014). Contextually variable signals can be functionally referential.

698   *Animal Behaviour,* 100: 1-8.

699   Schel AM, Machanda Z, Townsend SW, Zuberbühler K, & Slocombe KE (2013) Chimpanzee

700   food calls are directed at specific individuals. *Animal Behaviour* 86(5):955-965.

701   Schel AM, Townsend SW, Machanda Z, Zuberbühler K, & Slocombe KE (2013) Chimpanzee

702   alarm call production meets key criteria for intentionality. *PloS ONE* 8(10):e76674.

703   Segerdahl P, Fields W, & Savage-Rumbaugh ES. (2005). *Kanzi's primal language: The*

704   *cultural initiation of primates into language*.New York:Palgrave Macmillan.

705   Seyfarth RM, Cheney DL, & Marler P. (1980). Vervet monkey alarm calls: Semantic

706   communication in a free-ranging primate. *Animal Behaviour 28:* 1070-1094.

707   Seyfarth RM, Cheney DL, & Marler P. (1980). Monkey responses to three different alarm

708   calls: evidence for predator classification and semantic communication." *Science,*

709   *210:* 801-803.

710   Slocombe KE, Kaller T, Turman L, Townsend SW, Papworth S, Squibbs P,

711   Zuberbühler K. (2010). Production of food-associated calls in wild male

712   chimpanzees is dependent on the composition of the audience. *Behavioural Ecology*

713   *and Sociobiology* 64(12):1959-1966.

714   Slocombe KE, & Zuberbühler K. (2007). Chimpanzees modify recruitment screams as a

715   function of audience composition. *Proceedings for the National Academy of Sciences of*

716   *the United States of America*, *104*, 17228-17233.

717   Tabachnick BG & Fidell LS (2001) Using multivariate statistics. Boston, MA: Pearson.

718   Taglialatela JP, Reamer L, Schapiro SJ, & Hopkins WD. (2012). Social learning of a

719   communicative signal in captive chimpanzees. *Biology Letters, 8*(4), 498-501.

720   Taglialatela JP, & Savage-Rumbaugh S. (2003). Vocal production by a language-competent

721   *Pan paniscus*. *International Journal of Primatology, 24*(1), 1-17.

722   Tomasello M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.

723   Tomasello M, & Carpenter M. (2007). Shared intentionality. *Developmental Science, 10*(1):

724   121-125.

725   Tomasello M, Carpenter M, Call J, Behne T, & Moll H. (2005). Understanding and sharing

726   intentions: The origins of cultural cognition. *Behavioral & Brain Sciences, 28*, 1-17.

727   Townsend SW, Deschner T, & Zuberbuehler K (2008) Female chimpanzees use copulation

728   calls flexibly to prevent social competition. *Plos ONE* 3(6).

729   Van Krunkelsven, E., Dupain, J., Van Elsacker, L., & Verheyen, R. F. (1996). Food calling in

730   captive bonobos (*Pan paniscus*): an experiment. *International Journal of Primatology,*

731   *17*, 207-217.

732 Watson D, Wiese D, Vaidya J, & Tellegen A (1999) The two general activation systems of

733   affect: Structural findings, evolutionary considerations, and psychobiological

734   evidence. *Journal of Personality and Social Psychology,* 76(5):820.

735 Wich SA, Swartz KB, Hardus ME, Lameira AR, Stromberg E, & Shumaker RW. (2009). A case

736   of spontaneous acquisition of a human sound by an orangutan. *Primates, 50*(1), 56-

737   64.

738 Wilkinson GS, & Boughman WJ. (1998). Social calls coordinate foraging in greater spear-

739   nosed bats. *Animal Behaviour, 55*(2): 337-350.

740 Zuberbühler K. (2003). Referential signaling in non-human primates: Cognitive precursors

741   and limitations for the evolution of language. *Advances in the Study of Behavior: 33*,

742   265-307.

743 Zuberbühler K. (2006). Alarm calls - Evolutionary and cognitive mechanisms. In K. Brown

744   (Ed.), *Encyclopedia of Language and Linguistics* (2 ed., pp. 143-155). Oxford:

745   Elsevier.

746

747

748

Figure 1(on next page)

Figure 1

Figure 1. Time–frequency spectrograms illustrating peeps produced by four wild bonobos (BE, EM, ZD = male, NI = female) during different behavioural contexts. The emotional valence of the context is indicated in parantheses.
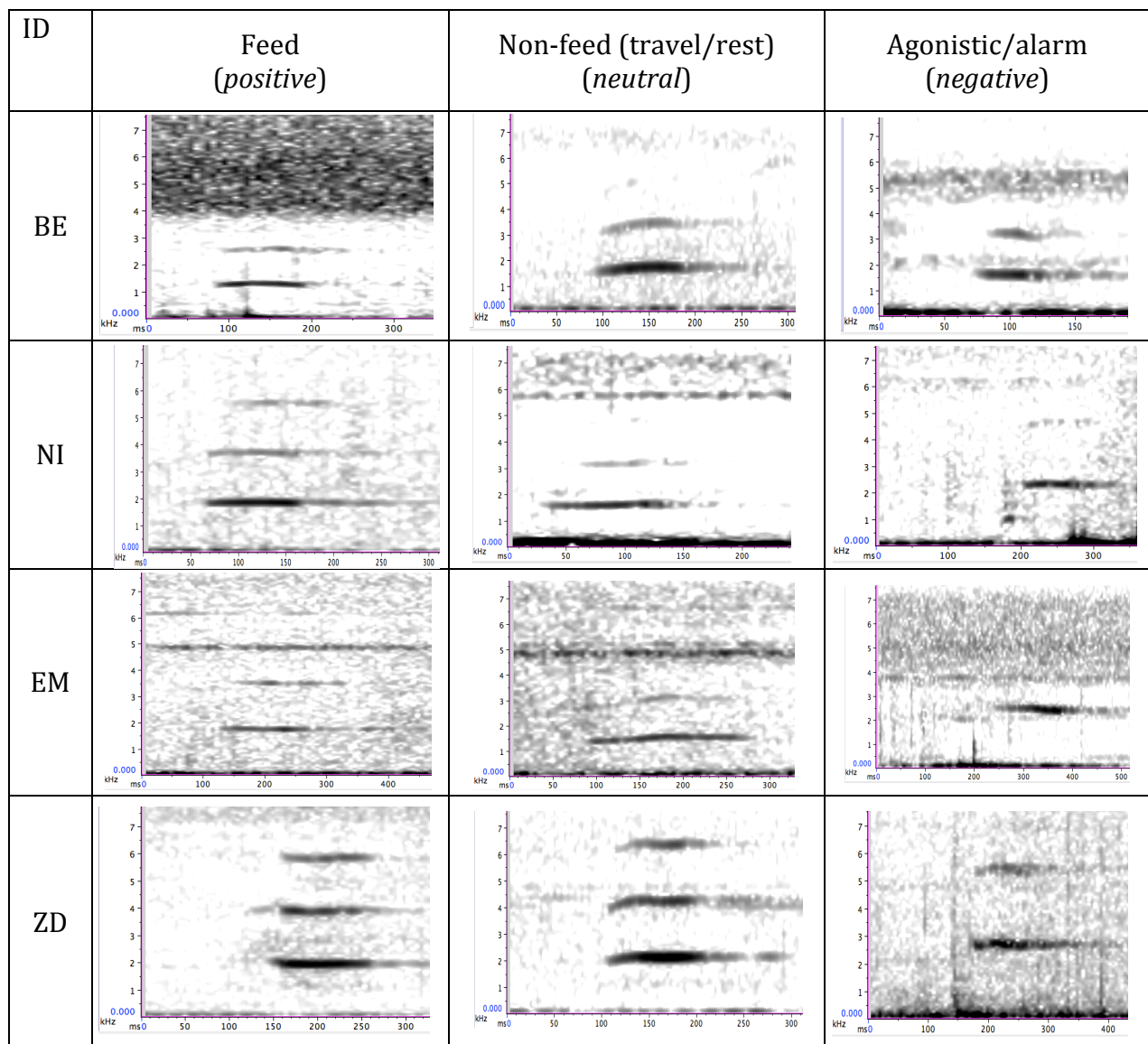
ID	Feed (*positive*)	Non-feed (travel/rest) (*neutral*)	Agonistic/alarm (*negative*)
BE			
NI			
EM			
ZD			

Figure 2(on next page)

Figure 2

Figure 2. Boxplots indicating six acoustic parameters of peep vocalisations that varied as a function of behavioural context. The emotional valence associated with the context is indicated in parentheses. Thick black lines represent medians; open circles and small asterisks represent outliers, box edges represent the upper and lower hinges of the H-spread, which generally matches the upper and lower quartiles; whiskers represent the adjacent values, which are the most extreme values still lying within hinges and the normal distribution of the sample. For significant differences, lines with ** represents $P < .05$, *** represents $P < .001$.