



HAL
open science

The reliability of individual vocal signature varies across the bonobo's graded repertoire

Sumir Keenan, Nicolas Mathevon, Jeroen M.G. Stevens, Florence Nicolè, Klaus Zuberbühler, Jean-Pascal Guéry, Florence Levréro

► To cite this version:

Sumir Keenan, Nicolas Mathevon, Jeroen M.G. Stevens, Florence Nicolè, Klaus Zuberbühler, et al.. The reliability of individual vocal signature varies across the bonobo's graded repertoire. *Animal Behaviour*, 2020, 169, pp.9 - 21. 10.1016/j.anbehav.2020.08.024 . hal-03492014

HAL Id: hal-03492014

<https://hal.science/hal-03492014>

Submitted on 17 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **The reliability of individual vocal signature varies across the bonobos'**
2 **graded repertoire**

3
4

5 Sumir Keenan^{1,2}, Nicolas Mathevon¹, Jeroen MG Stevens³, Florence Nicolè⁴,
6 Klaus Zuberbühler^{2,5}, Jean-Pascal Guéry⁶ and Florence Levréro¹

7
8

9 ¹University of Lyon / Saint-Etienne, Equipe Neuro-Ethologie Sensorielle, ENES
10 / CRNL, CNRS, INSERM, Saint-Etienne, 42100, France

11 ²University of St. Andrews, Department of Psychology & Neuroscience, St.
12 Andrews, Scotland, KY16 9JP, UK

13 ³APELAB - Behavioural Ecology and Ecophysiology Group, Department of
14 Biology, University of Antwerp, Belgium

15 ⁴University of Lyon / Saint-Etienne, BVPam Lab, CNRS FRE 3727, Saint-
16 Etienne, 42100, France

17 ⁵Université de Neuchâtel, Department of Comparative Cognition, Neuchâtel,
18 2000, Switzerland

19 ⁶Vallée des Singes Zoological Park, Romagne, 86700, France

20
21

22 **Corresponding authors:**

23 Sumir Keenan sumirkeenan@gmail.com

24 Florence Levréro florence.levrero@univ-st-etienne.fr

25
26

27 **Highlights**

- 28 • Coding of individual signature in a graded vocal repertoire
- 29 • Comparison of the strength of individual vocal signature across bonobo
30 repertoire.
- 31 • Quantitative analysis revealed a graded structure between the most
32 common bonobo call types.
- 33 • All call types investigated show significant individual vocal signatures.
- 34 • The individual signature is stronger in one extreme of the vocal
35 gradation.
- 36

37 **Abstract**

38

39 Animal vocalisations often contain both ‘dynamic’ information, related to short-
40 term fluctuations in the emitter’s emotional states, and ‘static’ information,
41 related to long-term attributes such as age, sex, weight and body size which
42 define the emitter’s “individual vocal signature”. While both types of information
43 may be of functional value to receivers, dynamic information requires acoustic
44 versatility, while static information depends on acoustic stability. Here we
45 investigate whether an individual vocal signature is present across the vocal
46 repertoire of the bonobo, *Pan paniscus*. We first emphasize the graded
47 character of the bonobo’s repertoire by describing the acoustic structure of its
48 five most common tonal vocalisations. We then evaluate the reliability of
49 identity information across these call types. The results show that, while all call
50 types support information related to the emitter’s identity, the reliability of these
51 vocal signatures was not consistent along the graded vocal continuum. Caller
52 identity is strongly encoded at one end of the acoustic gradation (high-hoot)
53 and decreases from bark, soft bark, peep-yelp to peep calls. Strikingly, the
54 reliability of the individual signature thus decreases from calls used in high-
55 arousal contexts to low-arousal contexts. To the best of our knowledge, this is
56 the first demonstration that an acoustic gradation that codes for ‘dynamic’
57 information can be accompanied by variation of the ‘static’ information that
58 supports vocal individuality.

59

60

61 **Keywords:** acoustic grading; bonobo; identity information; individual vocal
62 signature

63 INTRODUCTION

64

65

66 Bird and mammal vocalisations usually contain both ‘dynamic’ information,
67 related to short-term fluctuations in the emitter’s physiological and
68 psychological states, and ‘static’ information, related to idiosyncratic features
69 such as age, sex, weight, and body size which define the emitter’s “individual
70 vocal signature” (Briefer, 2012, 2020; Charlton, Pisanski, Raine & Reby, 2020;
71 Pisanski, Nowak, & Sorokowski, 2016; Raine, Pisanski & Reby, 2017; Taylor,
72 Charlton & Reby, 2016). Both dynamic and static information likely have
73 important roles in a wide range of social contexts (Taylor, Charlton & Reby,
74 2016). The acoustic coding of dynamic information relies on versatile acoustic
75 features that can be modulated depending on the emitter’s current internal
76 state (Briefer 2012, 2020; Pisanski, et al., 2016a; Raine, Pisanski & Reby,
77 2017; Taylor, Charlton & Reby, 2016), and can be reflected in a repertoire of
78 discreet call types (Bradbury & Vehrencamp, 1998; Cheney & Seyfarth, 2018;
79 Maynard-Smith & Harper, 2003;), and/or in more subtle acoustic variations
80 within each call type (graded vocalisations; Briefer, 2012, 2020). Sometimes,
81 the acoustic boundaries between call types are unclear that the whole vocal
82 repertoire can be considered as a graded system (e.g. Fischer, Wadewitz &
83 Hammerschmidt, 2017; Hammerschmidt & Fischer 1998; Keenan, Lemasson
84 & Zuberbühler, 2013; Manser et al., 2014; Marler 1977; Tallet et al., 2013).
85 While dynamic information requires acoustic versatility, static information
86 defining an “individual vocal signature” should depend on acoustic stability.
87 How do animal signals deal with this potential conflict of information coding?

88 One solution is temporal segregation, which has been demonstrated in the
89 vocalisations of the banded mongoose (*Mungos mungo*). This species
90 produces calls containing two successive distinct segments, the first being
91 stable and individually distinct, and the second being graded and correlating
92 with the emitter's current behaviour (Jansen, Cant & Manser, 2012). Whether
93 temporal segregation is a rare or a widespread way of separately coding
94 dynamic and static information in vocal signals is a question that remains to be
95 answered.

96 Another solution would be to keep the indexical acoustic cues supporting
97 individual signatures stable across the whole vocal repertoire. This would
98 mean a common set of individualized acoustic features, either shared between
99 calls when the repertoire is composed by discreet calls or remaining invariable
100 throughout a graded repertoire. Unfortunately, this hypothesis has been barely
101 tested, as the vast majority of studies investigating individual vocal signatures
102 have only focused on a single vocalisation type (e.g. in birds: blue footed
103 boobies, *Sula nebouxii*, Dentressangle, Aubin & Mathevon, 2012; jungle
104 crows, *Corvus macrorhynchos*, Kondo, Izawa & Watanabe, 2010; multiple
105 species of penguins, Aubin & Jouventin 2002 ; in mammals: bats, *Saccopteryx*
106 *bilineata* and *Noctilio albiventris*, Knörnschild & Von Helversen 2008; Voigt-
107 Heucke, Taborsky & Dechmann, 2010; elephants, *Loxodonta africana*, Soltis,
108 Leong & Savage, 2005; hyenas, *Crocuta crocuta*, Mathevon, Koralek,
109 Weldele, Glickman & Theunissen, 2010; marmots, *Marmotta sp.*, Matrosova,
110 Blumstein, Volodin & Volodina, 2011; seals, *Arctocephalus tropicalis*, Charrier,

111 Mathevon & Jouventin, 2003; northern elephant seals, *Mirounga*
112 *angustirostris*, Casey, Charrier, Mathevon & Reichmuth, 2015;
113 chimpanzees, *Pan troglodytes schweinfurthii*, Levréro & Mathevon 2013, and
114 spider monkeys, *Ateles geoffroyi*, Chapman & Weary 1990). To the best of our
115 knowledge, there has only been one study that has examined the coding of
116 individual signature throughout a whole repertoire composed by discreet calls,
117 in the zebra finch *Taeniopygia guttata* (Elie & Theunissen, 2018). In this
118 songbird, individual recognition is of primary importance for pair bonding
119 (Vignal, Mathevon & Mottin, 2004). Elie & Theunissen (2018) showed that
120 each zebra finch call type displays a distinct individual signature, and, contrary
121 to expectations, there is no common set of static, individualized, acoustic
122 features. Instead, the vocal repertoire of a given emitter supports numerous
123 signatures, and receivers have to memorize all of them to perform individual
124 recognition. Moreover, the Elie & Theunissen (2018) zebra finch study, as well
125 as the few other studies that have analysed more than one call type within a
126 species' vocal repertoire, have emphasised that the reliability of individual
127 signatures may vary among calls (e.g. putty-nosed monkeys, *Cercopithecus*
128 *nictitans*, Price, Arnold, Zuberbuhler & Semple 2009; fallow deer, *Dama dama*,
129 Vannoni & McElligott 2007, but see western gorilla, *Gorilla gorilla*, Salmi,
130 Hammerschmidt & Doran-Sheehy, 2014). For instance, Charrier, Jouventin,
131 Mathevon & Aubin (2001) found that both courtship and contact calls of a
132 marine bird, the South Polar Skua (*Catharacta maccormicki*), have a higher
133 potentiality of individual identity coding than the alarm call. In the red-capped

134 mangabey (*Cercocebus torquatus*), a monkey living in the African rainforest,
135 individual distinctiveness is higher in contact and threat calls than in other
136 vocalisations (Bouchet, Blois-Heulin, Pellier, Zuberbuhler & Lemasson, 2012).
137 Recently, Rubow, Cherry & Sharpe (2017) found similar results among the
138 repertoire of the dwarf mongoose (*Helogale parvula*).

139 While a proximal explanation for signature inconstancy within a repertoire
140 likely depends on variations in the shape of the vocal tract accompanying
141 modulation of acoustic production (review in Charlton, Pisanski, Raine & Reby,
142 2020), its functional value remains uncertain. Although it may have no specific
143 functional consequences (e.g. Rendall, Owren & Rodman, 1998), alternative
144 hypotheses state that individual signatures differ across call types as an
145 adaptive response to various socio-ecological pressures. On one hand, the
146 ‘distance communication hypothesis’ predicts that in low visibility
147 environments, such as dense forests, natural selection will favour individual
148 signatures in vocal signals used to communicate over long distances beyond
149 the visible range (Ey & Fischer 2009; Mitani & Stuht 1998). This hypothesis
150 has received some support by comparing two chimpanzees’ vocalisations, a
151 long distance and a short-range call (Mitani, Gros-Louis & Macedonia, 1996).
152 On the other hand, the ‘social function hypothesis’ (Price, Arnold, Zuberbuhler
153 & Semple, 2009; Snowdon & Hausberger, 1997) predicts that the reliability of
154 individual vocal signatures depends on the social role of a call type. If the
155 vocalisation is emitted in a context where being recognized is of primary
156 importance for individual-specific relationships, such as mated pair bonds,

157 mother-young bond, or tolerant relationships between neighbours in territorial
158 animals, it should bear reliable information about individual identity (Kondo &
159 Watanabe, 2009; Kreiman & Sidtis, 2011; Wiley, 2013). This hypothesis has
160 been supported in birds (Charrier, Jouventin, Mathevon & Aubin, 2001; Elie &
161 Theunissen, 2018) and some non-human primate species (e.g. mouse
162 lemurs, *Microcebus murinus* (Leliveld, Scheumann & Zimmermann, 2011),
163 Campbell's monkeys, *Cercopithecus campbelli campbelli* (Lemasson &
164 Hausberger, 2011). However, these distant-communication and social function
165 hypotheses may not be mutually exclusive. For instance, distance calls can
166 play a significant role in social regulation, facilitating cooperation, such as the
167 recruitment of conspecifics for support (Gersick, Cheney, Schneider, Seyfarth,
168 & Holekamp, 2015; Mitani & Nishida, 1993), coordinating travel (Schamberg,
169 Cheney, Clay, Hohmann & Seyfarth, 2016), signalling food ownership (Gros-
170 Louis 2004; Heinrich & Marzluff, 1991), or conveying the aggressive or
171 peaceful intentions of callers (Searcy, Anderson & Nowicki, 2006; Silk 2002).

172 To explore how an animal vocal repertoire can deal with the potential conflict
173 of coding both dynamic information (e.g. expressing motivation,
174 aggressiveness...) and static information (i.e. a vocal signature supporting
175 individual recognition), here we investigate variations in the strength of vocal
176 individual signature throughout the graded vocal repertoire of an ape species,
177 the bonobo. Bonobos display a complex vocal repertoire with highly graded
178 call types used flexibly across contexts (Bermejo & Omedes, 1999; de Waal,
179 1998). Despite growing interest in the acoustic communication of this species,

180 our current knowledge of the information content of bonobo calls is limited, as
181 only few studies have qualitatively described the bonobo vocal repertoire
182 (Clay, Archbold & Zuberbuhler, 2015; Clay & Zuberbuhler, 2009, 2011a,
183 2011b; Keenan et al., 2016; White, Waller, Boose, Merrill & Wood, 2015).

184 Bonobo society is characterised by co-dominance between the sexes
185 (Surbeck & Hohmann, 2013) and complex fission-fusion dynamics which
186 require sophisticated social knowledge and communication (Clay, Archbold &
187 Zuberbuhler, 2015; Furuichi, 2011). It has been suggested that individual vocal
188 recognition is essential for successful social navigation (White, Waller, Boose,
189 Merrill & Wood, 2015). In a recent work (Keenan et al., 2016), we
190 experimentally demonstrated that bonobos can vocally identify familiar to
191 unfamiliar individuals using the peep-yelp, a soft vocalisation used in short
192 range interactions. This study already revealed that soft calls convey
193 information about individual signature, which is extracted by conspecifics.
194 Other calls from bonobos' repertoire are good candidates to convey individual
195 signatures regarding their propagation capacity in close habitat and their
196 functions. Bonobos loud call types, e.g. high-hoots, can be heard from a
197 distance of 500 meters in the forest (Hohmann & Fruth, 1994), and are
198 regularly used by community members to communicate with one another when
199 the group splits into foraging parties throughout the day. These vocalisations
200 enable the sub-parties to convene around a resource, such as a fruit-bearing
201 tree or nesting tree, despite potential distances between parties (Bermejo &
202 Omedes, 1999; Hohmann & Fruth, 1994, 1995; Schamberg, Cheney, Clay,

203 Hohmann & Seyfarth, 2016, 2017; White, 1996; White, Waller, Boose, Merrill &
204 Wood, 2015). High-hoots are also exchanged when different communities
205 encounter one another, and they appear to influence inter-community
206 interactions (Furuchi, 2011; Hohmann & Fruth, 2002). Caller identity signalling
207 may thus be present in the bonobos' calls with various acoustic structure and
208 functions.

209 In this study, we test the hypothesis that an individual signature does exist
210 across the graded vocal repertoire of the bonobo, but that the reliability of this
211 signature varies along with the variation of the acoustic features defining the
212 call types. We firstly emphasize the graded character of the bonobo's
213 repertoire by providing a quantitative description of the acoustic structure of
214 the five most common tonal call types, as well as their contextual use. We then
215 compare the individual distinctiveness across call types.

216

217 **METHODS**

218

219 **Ethics Statement:**

220 All research conducted for this research paper was observational and no
221 experimental manipulations occurred. All data collection protocols were
222 performed in accordance with the relevant guidelines and regulations, and
223 were approved by the Institutional Animal Ethical Committee of the University
224 of Lyon/Saint-Etienne, under the authorization no. 42-218-0901-38 SV 09 (Lab
225 ENES).

226

227 **Subjects**

228 The bonobos observed for this study were members of three separate captive
229 groups housed at three European zoos: Apenheul (Apeldoorn, the
230 Netherlands), Planckendael Zoo (Mechelen, Belgium) and la Vallée des
231 Singes (Romagne, France; see Supplemental Table A1 for group composition
232 at each zoo). At each zoo, groups were housed in large indoor enclosures with
233 varying access to off-exhibit rooms and outdoor islands. All individuals
234 included in the study had lived in similar captive zoo environments certified by
235 the European Association of Zoos and Aquaria for a minimum of 10 years prior
236 to recording. Vocal recordings and observations were taken from all areas at
237 all three zoos, except in the off-exhibit enclosures at Apenheul.

238

239 To avoid the potential confound of differences in vocal tract size due to age,
240 only bonobos over the age of 10 were included in this study. Additionally, one
241 adult male and one female, who had overall low calling rates, were excluded.
242 This led to a total of 21 individuals ranging in age from 10 – 45, with a mean
243 age of 20.5 years old, and comprised of 13 females and 8 males.

244

245 **Data Collection**

246 *Vocal recordings*

247 Vocal recordings were collected between March 2013 and March 2014,
248 beginning no earlier than 8 a.m. and finishing no later than 6 p.m. Recordings

249 at Apenheul Zoo were collected from May 14, 2013 – July 6, 2013, as well as
250 on March 12, 2014, and amounted to 175 hours of total recording time.
251 Recordings at Planckendael Zoo were collected from March 20, 2013 – May
252 10, 2013 and from February 20 – March 4, 2014, and amounted to 190 hours
253 of total recording time. Recordings at la Vallée des Singes were collected from
254 October 28, 2013 – November 25, 2013, and amounted to 115 hours of total
255 recording time.

256

257 Audio recordings were taken using a Zoom H4 Digital Multi-track Recorder
258 (44.1 kHz sample rate, 16 bits per sample, .wav files) - recording in stereo,
259 with one channel devoted to a Sennheiser MKH70-1 ultra-directional
260 microphone recording any bonobo vocal behaviour and the second channel
261 connected to a micro-tie recording device, model AKG MPA III, for comments
262 by the researcher. This allowed for temporal synchronising of each
263 vocalisation to information on vocaliser identity and call context, as recorded
264 by the researcher.

265

266 *Call types*

267 Each call was assigned to one of the 11 types based on classifications
268 described in previous studies (Bermejo & Omedes, 1999; Clay & Zuberbuhler,
269 2009; de Waal, 1988) by SK. We calculated the frequency with which each call
270 type was produce so that we could assign a proportion to each call type in the
271 entire dataset (relative to the 2,373 measurable calls collected from the 21

272 adults included in the study). Any call type that represented less than 10% of
273 the total number of measurable calls collected was excluded from the analysis.
274 In addition, two call types, screams and grunts, were excluded from the current
275 study, due to their noisy acoustic features characterised by non-linear
276 phenomena and lack of clear harmonic structure. Unlike the rest of the adult
277 vocal repertoire, these two call types would require a different set of
278 measurements to describe their acoustic features. In total, five call types were
279 retained, which represented the vast majority of calls emitted by all individuals
280 in the study groups (78% of all calls collected, excluding screams and grunts):
281 (1) high-hoots (14%; described as 'staccato-high hoots' by de Waal (1988) and
282 'high hoots' by Bermejo & Omedes (1999)); (2) barks (18%; described as
283 'barks' and 'wieew barks' by de Waal (1988) and as 'barks' and 'composed
284 barks' by Bermejo & Omedes (1999)); (3) soft barks (17%; described as 'soft
285 barks' by Bermejo & Omedes (1999), 'food barks' by Clay & Zuberbuhler,
286 (2009) and not described by de Waal (1988); (4) peep-yelps (18%; described
287 as 'peep-yelps' by de Waal (1988); Bermejo & Omedes (1999); Clay &
288 Zuberbuhler (2009) and (5) peeps (11%; described as 'peeps' by de Waal
289 (1988); Bermejo & Omedes (1999); Clay & Zuberbuhler (2009); see
290 Supplemental Table A2 for full acoustic description of each call type).

291

292 *Call contexts*

293 If a caller could be identified, social and individual contextual information was
294 also recorded. Each recorded call occurred in one of the following eight

295 contexts: 1) 'pre-feeding' – calls given directly prior to or at the start of
296 scheduled feedings when group/social excitement was extremely high
297 (excitement described as increased sexual activity, displacements, pacing,
298 pilo-erection, vocal activity, displays and increased likelihood of aggression);
299 2) 'feeding' – calls given during scheduled feedings; 3) 'foraging' – calls given
300 when foraging for or eating food found outside or in the inside enclosure
301 outside of scheduled feeding times; 4) 'aggression' - calls given during
302 agonistic encounters, including mild aggression (no physical contact),
303 aggression (mild physical contact, such as hitting, kicking or grabbing between
304 only two individuals), conflict (a range of physical contact, including biting, and
305 often between more than two individuals) - agonistic interactions were also
306 classified into calls from victims, from aggressors or from bystanders; 5)
307 grooming - calls given during bouts of grooming; 6) contact – calls given when
308 a subject was resting or moving but nothing else; 7) change of environment –
309 calls given when individuals were shifted from one enclosure to another or
310 from inside to outside; and 8) external event – calls given when an individual
311 was visibly startled by or responding to a disturbance external to the group.

312

313 **Acoustic Analysis**

314

315 Only vocalisations that could unequivocally be assigned to one caller were
316 retained for analysis. Only calls of good, measurable quality were included,
317 while calls that overlapped with background noise (e.g. birds chirping, water

318 falling, zoo visitors speaking or other bonobos calling) were removed from the
319 dataset. In total, 1,850 individual calls were retained for analyses (with an
320 average of 88 total calls per individual ($N=21$), S.D.= 37.87, minimum=45,
321 maximum=227). Raven Pro 1.3 was used to measure automatic and manual
322 parameters on each call and spectrograms were generated using a 512-
323 sample Hann window (50% frame overlap, frequency resolution of 86.1 Hz and
324 temporal resolution of 11.6 mS). A correlation matrix was produced and very
325 highly correlated variables were removed (0.9 and above), resulting in a total
326 of 16 measurements being retained for analysis – nine manually measured
327 parameters describing the fundamental frequency and its temporal modulation,
328 and seven automatically computed parameters describing the distribution of
329 energy across the frequency spectrum of the entire call (Table 1; Figure 1).

330

331

332 **Statistical Analysis**

333

334 *Call type distinctiveness*

335 To confirm whether our dataset could be reliably classified into the five main
336 call types proposed by previous studies, we performed a multivariate permuted
337 discriminant function analysis (pDFA) to determine the statistical distinction
338 between the 1,850 individual calls (high-hoots: $n=333$, barks: $n=431$, soft
339 barks: $n=413$, peep-yelp: $n=420$, peeps: $n=253$; See Supplemental Table A2
340 for acoustic description of each call type). The raw values of the 16 acoustical

341 parameters of interest were centred and normalised by transforming them into
342 z-scores as the parameter set consisted of different units.

343

344 Mundry and Sommer (2007) have convincingly argued that using traditional
345 discriminant function analyses (DFA), when analysing non-independent data
346 (e.g. if the same individual contributed multiple calls), is a case of pseudo-
347 replication and can inflate results. To address this issue, we employed a
348 permuted discriminant function analysis (pDFA - Mathevon et al., 2010;
349 Mundry & Sommer 2007) using the 16 acoustic variables (analysis performed
350 in R, Version 3.2.0). The pDFA takes the classic DFA a step further by
351 comparing the distribution of percent correct classifications obtained after 100
352 iterations (mean effect size) to the distribution of percent correct values
353 obtained by initially randomly assigning the call type to each individual call (for
354 the current analysis this distribution was obtained via 1000 created data sets
355 where the call type of each call was randomly permuted).

356

357 In detail, we performed the following steps. In the first step of the DFA, a
358 training data set was used to generate a set of linear discriminant functions.
359 The training data set consisted of randomly selected sounds from each
360 individual. The number of sounds selected per individual was the same for all
361 individuals and equal to 2/3 of the smallest number of sounds that we obtained
362 for an animal in our data set. In the second step, the discriminant functions
363 generated from the training data set were used to classify the remaining

364 sounds. For each individual, at least 1/3 of the sound provided by each
365 individual was thus included in the validating data set. This cross-validation
366 step gives a measure of the effect size (the percentage of correctly classified
367 sounds; which has to be compared with chance, here 20%, i.e. 1/5 possible
368 call types). We ran 100 iterations of these two-step DFAs with both training
369 and validation data sets chosen at random. The mean effect size (mean
370 percentage of correctly classified sounds) was obtained by calculating the
371 average of the percentages of correctly classified sounds obtained with each
372 of the 100 validation data sets. In addition to the cross-validated DFAs
373 performed on original data sets, new data sets were also created, where the
374 identity of sounds was randomly permuted between individuals (permuted
375 DFA), to obtain the statistical significance of the mean effect size. From these
376 randomised sets, the same steps, fitting and validation, were consecutively
377 performed. After 1000 iterations, we calculated the proportion of randomized
378 validation data sets revealing a number of correctly classified calls being at
379 least as large as the effect size obtained with the non-randomised validation
380 data set. This proportion gives the significance of the discrimination level and
381 is equivalent to a p-value [Dentressangle, Aubin & Mathevon, 2012; Mathevon,
382 Koralek, Weldele, Glickman & Theunissen, 2010; Mundry & Sommer 2007].

383

384 *Individual vocal signatures*

385 We assessed the reliability of individual signatures for each call type
386 independently using two different approaches: the first approach used a pDFA

387 (same method as described above, except that calls within each call type were
388 classified according to the identity of the vocaliser; here the chance of correct
389 classification was 10%, i.e. 1/10 possible individual callers), and secondly by
390 calculating the acoustic variation and potential for individual coding (PIC) for
391 each call type overall, as well as for each acoustic parameter describing the
392 calls (Robisson, Aubin & Bremond, 1993).

393

394 The amount of variability across the five call types was determined by
395 calculating the inter- and intra-individual coefficients of variation (CV). Intra-
396 individual CVs correspond to the variability of each acoustic variable within
397 individuals, and inter-individual CVs correspond to the variability of each
398 acoustic variable between individuals (see mathematical formula below). The
399 CV values were then used to calculate potentials of individual coding (PICs)
400 for each acoustic variable in each call type (according to Robisson, Aubin &
401 Bremond, 1993). For each considered acoustic parameter, the PIC
402 corresponds to the ratio between the variability between individuals and the
403 mean variability within individuals. A PIC value greater than one suggests that
404 the acoustic parameter considered may be used for individual recognition as
405 its intra-individual variability is smaller than its interindividual variability.
406 Additionally, recent studies investigating individual distinctiveness in non-
407 human primates have calculated CV and PIC values, including them here
408 enables direct comparison between species (Bouchet, Blois-Heulin, Pellier,
409 Zuberbuhler & Lemasson, 2012; Lemasson & Hausberger, 2011; Salmi,

410 Hammerschmidt & Doran-Sheehy, 2014). As the coefficient of variation (CV)
411 can only be calculated with variables on a ratio scale, two variables (that were
412 on an interval scale) were not included (Slope– F₀ Start to Mid and Slope– F₀
413 Mid to End). Therefore, for each of the remaining 14 parameters in each call
414 type separately we first calculated the $CV(\text{Inter}) = (100 \times \text{S.D. calculated}$
415 $\text{across all individuals}) / (\text{Mean (calculated across all individuals)})$ and the
416 $CV(\text{Intra}) = \text{the mean of individual CV values, where the CV (= } 100 \times$
417 $\text{S.D./Mean})$ was calculated for each individual separately. PIC values could
418 then be obtained for each parameter ($\text{PIC} = \text{CV}(\text{Inter}) / \text{CV}(\text{Intra})$). To assess
419 the variation of the five call types as a whole, we took the mean of the
420 $CV(\text{Inter})$ and the $CV(\text{Intra})$ across the 14 parameters for each call type
421 separately. These means were then used to calculate the PIC values for each
422 call type. To test for call type differences in the levels of variation and potential
423 for individual coding, Friedman tests were done with the CV and PIC values for
424 each parameter for each call type respectively.

425

426 For the acoustic variation and PIC analyses, we included all individuals with
427 12-20 calls for each call type. For individuals who had more than 20 calls in a
428 given call type, we randomly chose calls to be excluded (however ensuring
429 where possible that a single calling event or calls given in a single day were
430 not over-represented). This ensured that no single individual was
431 overrepresented, as no individual contributed more than 20 calls for each call
432 type.

433

434 For the pDFA analysis, we retained individuals who had a minimum of 14 calls
435 for each call type (with the exception of one individual who had 13 calls in the
436 peep call type) (See Supplemental Table A1 for the number of calls
437 contributed by each individual for each call type). Not every individual had a
438 sufficient number of calls for each call type to be included in all five-call type
439 analyses. The peep call-type had the fewest number of contributing individuals
440 ($n=10$). To allow for direct comparison across call types, we randomly chose
441 ten individuals for each of the other four call types (balancing the data for
442 individual sex and group), with the majority of individuals contributing to 2 or 3
443 call type analyses (mean = 2.5, max = 4, min = 1; Supplemental Table A1).

444

445 We considered the identity of callers to assess the individual vocal signature,
446 but not the sex, rank or age. Indeed, sex information is one of the many other
447 elements that constitute individual identity. A receiver may identify the overall
448 identity of a familiar caller and not need to categorize it by its sex. To identify
449 other parameters and their interactions that may influence the identity
450 signature was beyond the scope of this study.

451

452 **RESULTS**

453

454 **Call types – Classification and context**

455 The results of the pDFA showed that the acoustic structure of calls supports
456 the initial, human-driven, classification into call types, with an accuracy of 57%
457 (chance = 20%, $P < 0.001$; Figure 2).

458

459 A scatterplot using discriminant function 1 and 2 to visualise the grouping of
460 the call types, fails to show any distinct boundaries between the five call types,
461 suggesting acoustic grading (Figure 3). We observed a crescent shape from
462 the soft peep to loud high-hoot, which highlights the strong graded nature of
463 the most common call types in the vocal system of the bonobo.

464

465 Despite the gradedness of the five call types, overall the call types were used
466 significantly differently from one another in different contexts ($\chi^2 = 930.281$, P
467 < 0.001 ; results of Chi-square on contingency table of observed vs. expected
468 number of calls emitted in each context for each call type) (Table 2), and post-
469 hoc tests revealed that each call type was used significantly differently from
470 each other call type (see Supplemental Table A3 for details). The high-hoots
471 were used most often before feeding times (pre-feeding) and by aggressors in
472 agonistic encounters (aggression), but also during changes in environment
473 and to external events (or alarm). Barks were given in similar contexts to high-
474 hoots, however, with a marked increase of calls being given during feeding
475 events. Soft barks were given mostly in association with feeding and foraging,
476 but also during pre-feeding events. Peep-yelps and peeps, finally, were used
477 similarly with the main difference being an increased usage of peeps during

478 grooming and contact and a decreased usage during feeding or foraging
479 (Table 2).

480

481 **Individual vocal signature**

482 Five separate pDFA analyses were run on each call type to evaluate the level
483 of individual distinctiveness along the graded repertoire. Calls were assigned
484 to the correct emitter 53% of the time for high-hoots (chance level = 10%; $P =$
485 0.001), 44% for barks (chance level = 10%; $P = 0.001$), 30% for soft barks
486 (chance level = 10%; $P = 0.001$), 25% for peep-yelps (chance level = 10%; $P =$
487 0.004) and 23% for peeps (chance level = 10%; $P = 0.006$) (all percentages
488 were from cross-validated data sets; Figures 3 and 4; Supplemental Tables
489 A4, A5, A6, A7, A8).

490

491 **Variability and vocal signature**

492 All of the call types had PIC values over 1, indicating that all five have some
493 capacity to code for individuality, however the PIC was significantly stronger in
494 some call types (Friedman test of PIC values: $\chi^2 (13) = 49.114$, $P < 0.001$).
495 PIC was highest in high-hoots (1.22), followed by barks (1.18), peep-yelps
496 (1.10), soft barks (1.08) and peeps (1.03) (Table 3).

497

498 We then used this information to investigate which of the acoustic parameters
499 had the highest potential for individual coding. Across all call types the onset
500 frequency of the fundamental frequency (F_0 -Start) (1.20) and the ascending

501 slope (1.22) had the highest PIC values (Table 3). However, when
502 investigating each call type separately, these two parameters did not always
503 have the highest potential for individual coding. The onset frequency of the
504 fundamental frequency (F_0 -Start) had the highest PIC in both the peep and
505 barks, while the maximum frequency reached on the fundamental frequency
506 (F_0 -Peak) had the highest PIC in peep-yelps and high-hoots and the call
507 duration and the frequency of the fundamental at the end of the call (F_0 -End)
508 equally had the highest PIC for soft barks. Additionally, for each call type
509 separately not all acoustic parameters reached the minimum threshold (PIC \leq
510 1) of identity coding (Table 3). It is clear that acoustic parameters related to the
511 tonality had consistently higher PIC values than the parameters describing the
512 energy distribution for all five call types.

513

514 **DISCUSSION**

515

516 Our results demonstrate that the bonobo tonal calls investigated show clear
517 individual vocal signatures. Furthermore, the results support our hypothesis
518 that the reliability of this signature varies along the graded vocal continuum
519 formed by these calls. Caller identity signalling is stronger at one end of the
520 acoustic gradation (high-hoot) and then decreases from bark, soft bark, peep-
521 yelp to peep calls. The reliability of the individual signature thus decreases
522 from calls mostly used in high-arousal contexts to the ones used in low-arousal
523 contexts. To the best of our knowledge, this is the first demonstration that the

524 acoustic gradation that codes for 'dynamic' information can be accompanied
525 by a gradation of the strength of the 'static' information that supports vocal
526 individuality.

527

528 Individual signatures are the result of idiosyncratic acoustic features
529 characterizing each individual. These features are constrained by the
530 biomechanics of sound production, which varies depending on individual
531 morphology, anatomy, and physiology. As such, a proximal explanation for the
532 variation of the strength of individual signature across the vocal repertoire is
533 that the acoustic features of calls do not have the same potential for encoding
534 individual variation. Soft calls, namely characterized by little to no frequency
535 modulation and short durations, may offer reduced possibility for encoding
536 identity cues. Conversely, highly frequency modulated calls as the high-hoot
537 may support higher possibilities for individual differences.

538

539 Importantly, we recently demonstrated with playback experiments that
540 bonobos are able to discriminate between the peep-yelps of familiar and
541 unfamiliar individuals (Keenan et al., 2016). This suggests that even the call
542 types with a less reliable individual signature may contain enough identity
543 information for individual recognition in bonobos. Similar findings with various
544 species have demonstrated that individuals are able to utilize acoustically
545 encoded identity information to recognise others even in calls with low
546 potential for individuality (e.g. in zebra finches – Elie & Theunissen, 2018;

547 northern fur seals, *Callorhinus ursinus* – Insley, 2000; domestic horses, *Equus*
548 *caballus* – Proops, McComb & Reby, 2009; and rhesus macaques *Macaca*
549 *mulatta* – Rendell, Rodman & Emond, 1996). However, peep-yelps, and other
550 soft calls, may greatly suffer from long range propagation through the forest
551 environment. We assume that the individual signature carried by these calls
552 should be efficient only at short-range. We suggest that the increase in
553 signature reliability from peeps to high-hoots may be functionally relevant
554 since it allows louder calls, which propagate further, to maintain the fidelity of
555 individual information (“distance hypothesis”, Ey & Fischer, 2009; Fedurek,
556 Zuberbühler & Dahl, 2016; Mitani, Gros-Louis & Macedonia, 1996; Mitani &
557 Stuht 1998). It is also possible that the social value of individual recognition
558 increases from peeps to high-hoots (“social function hypothesis”, Price, Arnold,
559 Zuberbuhler & Semple, 2009; Snowdon & Hausberger, 1997; Snowdon &
560 Hausberger 1997).

561

562 Whether variation in the reliability of individual signatures across a vocal
563 repertoire is an adaptive response to socio-ecological pressures is still
564 debated. Similar to our data, a recent study (Salmi, Hammerschmidt & Doran-
565 Sheehy, 2014) investigating individual distinctiveness in both close and long
566 range call types of female western gorillas (*Gorilla gorilla gorilla*), found that all
567 call types had the same potential for individual coding. The authors concluded
568 that neither the social function nor the distance hypotheses accurately

569 reflected the possible evolutionary pressures likely acting on female gorilla
570 calls.

571

572 While gorillas are the only other ape species where multiple call types were
573 investigated, Mitani et al. (1996) studied two call types in chimpanzees, the
574 bonobo's congener. They found, as in our study, that loud calls contained
575 stronger identity information than soft calls used when in close contact with
576 other individuals. Chimpanzees and bonobos both live in dense rain forests,
577 forming far more complex fluid fission-fusion societies compared to the
578 gorilla's largely stable polygynous group structure (Robbins, Bermejo,
579 Cipolletta, Magliocca, Parnell, & Stokes, 2004). This difference in social
580 structure, and the need to regularly communicate with group members over
581 long distances, is likely a driving evolutive pressure for accurate identity
582 signaling over long distance.

583

584 So far acoustically graded calls have been overlooked in animal
585 communication research, despite being widespread in the vocal repertoire of
586 terrestrial mammals (Crockford, 2019; Fischer, Wadewitz & Hammerschmidt,
587 2017). This lack of knowledge is mainly due to the complexity of capturing and
588 describing the variations in the signal repertoires without forcing arbitrary
589 categorizations (Fischer, Wadewitz & Hammerschmidt, 2017). Yet, a
590 comparative perspective on graded repertoires of non-human primates and
591 other mammals may inform understanding of our own vocal repertoire. Indeed,

592 the non-verbal vocalisations of humans (e.g. laughter, cries, and screams)
593 resemble those of non-human mammals, especially when one considers how
594 they greatly fluctuate in their acoustic structure (Briefer 2012, 2020; Morton,
595 1977). Despite their importance in human social regulation, how individual
596 vocal signatures change or are preserved across calls used in a diversity of
597 communication contexts -from joy to distress- is also poorly known. A recent
598 study in humans revealed that individual differences in fundamental frequency
599 (F0), which is an important biosocial marker during speech production (for
600 reviews Lavan, Burton, Scott & McGettigan, 2019; Pisanski, Cartei,
601 McGettigan, Raine & Reby, 2016), may be preserved across non-verbal
602 sounds (from laughter to screams) in largely valence-specific manner. For
603 instance, individual differences in F0 were preserved across pain vocalisations
604 representing varying levels of pain intensity (Pisanski, Raine & Reby, 2020).
605 Pisanski et al. (2016b) argued that the high ability of humans to voice
606 modulation is likely to predate our ability to articulate the verbal dimension of
607 speech. Here we raise the question whether a vocally graded repertoire is to
608 some extent a form of vocal flexibility.

609

610

611 **Acknowledgments:**

612 We would like to thank the French Ministère de l'Enseignement Supérieur et
613 de la Recherche (PhD grant to SK), the Université de Saint-Etienne (research
614 sabbaticals to FL and NM, visiting professorship to KZ and research funding),

615 the European Research Council (KZ grant PRILANG 283871), Institut
616 universitaire de France (NM), and the Labex CeLyA. We warmly thank Colette
617 Bouchut and Nicolas Boyer for their endless support, Roger Mundry for the
618 pDFA method, and the zoological parks of Apenheul, Planckendael and La
619 Vallée des Singes for their welcome, and especially to the bonobo keepers for
620 their support and patience. Lastly, we thank the two anonymous reviewers
621 who have accepted to review this paper and have contributed to improve it.

622

623

624 REFERENCES

625

626 Aubin, T. & Jouventin, P. (2002). How to identify vocally a kin in a crowd? The
627 penguin model. *Advances in the Study of Behavior*, 31, 243-277. doi:
628 10.1016/S0065-3454(02)80010-9

629

630 Bermejo, M. & Omedes, A. (1999). Preliminary vocal repertoire and vocal
631 communication of wild bonobos (*Pan paniscus*) at Lilungu (Democratic
632 Republic of Congo). *Folia Primatologica*, 70, 328-357. doi:10.1159/000021717

633

634 Bouchet, H., Blois-Heulin, C., Pellier, A., Zuberbuhler, K. & Lemasson, A.
635 (2012). Acoustic variability and individual distinctiveness in the vocal repertoire
636 of red-capped mangabeys (*Cercocebus torquatus*). *Journal of Comparative*
637 *Psychology*, 126, 45-56. doi:10.1037/a0025018

638

639 Bradbury, J.W. & Vehrencamp, S.L. (1998). *Principles of Animal*
640 *Communication*. New York: Sinauer Associates.

641

642 Briefer, E. F. (2012). Vocal expression of emotions in mammals: mechanisms
643 of production and evidence. *Journal of Zoology*, 288(1), 1-20.

644

645 Briefer, E. F. (2020). Coding for 'Dynamic' Information: Vocal Expression of
646 Emotional Arousal and Valence in Non-human Animals. In N. Mathevon & T.
647 Aubin (Eds), *Coding Strategies in Vertebrate Acoustic Communication* (pp.
648 137-162). Springer, Cham.

649

650 Casey, C., Charrier, I., Mathevon, N., Reichmuth, C. (2015). Rival assessment
651 among elephant seals: evidence of associative learning during male-male
652 contests. *Royal Society Open Science*, 2, 150228.

653 doi.org/10.1098/rsos.150228

654

655 Charrier, I., Jouventin, P., Mathevon, N. & Aubin, T. (2001). Individual identity
656 coding depends on call type in the South Polar skua, *Catharacta maccormicki*.
657 *Polar Biology*, 24, 378-382. doi:10.1007/s003000100231

658

659 Charrier, I., Mathevon, N. & Jouventin, P. (2003). Vocal signature recognition
660 of mothers by fur seal pups. *Animal Behaviour*, 65, 543-550. doi:
661 10.1006/anbe/2003.2073

662

663 Chapman, C.A. & Weary, D.M. (1990). Variability in spider monkeys'
664 vocalisations may provide basis for individual recognition. *American Journal of*
665 *Primatology*, 22, 279-284. doi:10.1002/ajp.1350220407

666

667 Charlton, B. D., Pisanski, K., Raine, J., & Reby, D. (2020). Coding of Static
668 Information in Terrestrial Mammal Vocal Signals. In N. Mathevon & T. Aubin
669 (Eds), *Coding Strategies in Vertebrate Acoustic Communication* (pp. 115-136).
670 Springer, Cham.

671

672 Cheney, D. L. & Seyfarth, R. M. (2018). Flexible usage and social function in
673 primate vocalisations. *Proceedings of the National Academy of Sciences*, 115,
674 1974–1979. doi: 10.1073/pnas.1717572115

675

676 Clay, Z. & Zuberbuhler, K. (2009). Food-associated calling sequences in
677 bonobos. *Animal Behaviour*, 77, 1387-1396.
678 doi:10.1016/j.anbehav.2009.02.016

679

680 Clay, Z. & Zuberbuhler, K. (2011a). Bonobos extract meaning from call
681 sequences. *PLOS One*, 6, e18786. doi:10.1371/journal.pone.0018786

682

683 Clay, Z., & Zuberbuehler, K. (2011b). The Structure of bonobo copulation calls
684 during reproductive and non-reproductive sex. *Ethology*, *117*, 1158-1169. doi:
685 10.1111/j.1439-0310.2011.01975.

686 Clay, Z., Archbold, J. & Zuberbuehler, K. (2015). Functional flexibility in wild
687 bonobo vocal behaviour. *PeerJ*, *3*, e1124. doi: 10.7717/peerj.1124

688

689 Crockford, C. (2019). Why does the chimpanzee vocal repertoire remain
690 poorly understood?-and what can be done about it. In *The chimpanzees of the*
691 *Tai forest: 40 years of research* (pp. 394-409). Cambridge University Press.

692

693 Dentressangle, F., Aubin, T, & Mathevon, N (2012). Males use time whereas
694 females prefer harmony: individual call recognition in the dimorphic blue-footed
695 booby. *Animal Behaviour*, *84*, 413-420. doi: 10.1016/j.anbehav/2012/05/012

696

697 de Waal, F.B.M. (1988). The communicative repertoire of captive bonobos
698 (*Pan paniscus*) compared to that of chimpanzees. *Behaviour*, *106*, 183-251.
699 doi: 10.1163/156853988X00269

700

701 Elie, J.E. & Theunissen, F.E. (2018). Zebra finches identify individuals using
702 vocal signatures unique to each call type. *Nature Communications*, *9*, 4026.

703

704 Ey, E. & Fischer, J. (2009). The “acoustic adaptation hypothesis” – A review of
705 the evidence from birds, anurans and mammals. *Bioacoustics*, 19, 21-48.
706

707 Fedurek, P., Zuberbühler, K., & Dahl, C. D. (2016). Sequential information in a
708 great ape utterance. *Scientific Reports*, 6, 38226.
709

710 Fischer, J., Wadewitz, P., & Hammerschmidt, K. (2017). Structural variability
711 and communicative complexity in acoustic communication. *Animal Behaviour*,
712 134, 229-237.
713

714 Furuchi, T. (2011). Female contributions to the peaceful nature of bonobo
715 society. *Evolutionary Anthropology*, 20, 131-142. doi: 10.1002/evan.20308
716

717 Gersick, A.S., Cheney, D.L., Schneider, J.M., Seyfarth, R.M., & Holekamp,
718 K.E. (2015). Long-distance communication facilitates cooperation among wild
719 spotted hyaenas, *Crocuta crocuta*. *Animal Behaviour*, 103, 107-116.
720

721 Gros-Louis, J. (2004). The function of food-associated calls in white-faced
722 capuchin monkeys, *Cebus capucinus*, from the perspective of the signaller.
723 *Animal Behaviour*, 67(3), 431-440.
724

725 Hammerschmidt, K., & Fischer, J. (1998). The vocal repertoire of Barbary
726 macaques: a quantitative analysis of a graded signal system. *Ethology*, 104,
727 203-216.
728

729 Heinrich, B., & Marzluff, J. M. (1991). Do common ravens yell because they
730 want to attract others? *Behavioral Ecology and Sociobiology*, 28(1), 13-21.
731

732 Hohmann, G. & Fruth, B. (1994). Structure and use of distance calls in wild
733 bonobos (*Pan paniscus*). *International Journal of Primatology* 15, 767–782.
734 doi:10.1007/BF02737430
735

736 Hohmann, G., & Fruth, B. (1995). Loud calls in great apes: sex differences and
737 social correlates. In: E. Zimmermann (Ed.), *Current Topics in Primate Vocal*
738 *Communication* (pp.161–184). New York, New York: Plenum Press.
739

740 Hohmann, G. & Fruth, B. (2002). Dynamics is social organization of bonobos
741 (*Pan paniscus*). In: Boesch, C., Hohmann, G. & Marchant, L.F. (Eds.),
742 *Behavioural Diversity in Chimpanzees and Bonobos* (pp.138-150). Cambridge,
743 U.K.: Cambridge University Press.
744

745 Insley, S. J. (2000). Long-term vocal recognition in the northern fur seal.
746 *Nature*, 406, 404–5. doi:10.1038/35019064
747

748 Jansen, D.A., Cant, M.A. & Manser, M.B. (2012). Segmental concatenation of
749 individual signatures and context cues in banded mongoose (*Mungos mungo*)
750 close calls. *BMC Biology* 10, 97. doi:10.1186/1741-7007-10-97
751

752 Keenan, S., Lemasson, A., Zuberbühler, K. (2013). Graded or discrete? A
753 quantitative analysis of Campbell's monkey alarm calls. *Animal Behaviour*, 85,
754 109–118.
755

756 Keenan, S., Mathevon, N., Stevens, J.M.G., Guery, J.P., Zuberbuhler, K. &
757 Levrero, F. (2016). Enduring voice recognition in bonobos. *Scientific Reports*,
758 6, 22046. doi: 10.1038/srep22046.
759

760 Knörnschild, M., and Von Helversen, O. (2008). "Nonmutual vocal mother- pup
761 recognition in the greater sac-winged bat," *Animal Behaviour*, 76, 1001-1009.
762 <https://doi.org/10.1016/j.anbehav.2008.05.018>
763

764 Kondo, N. & Watanabe, S. (2009). Contact calls: information and social
765 function. *Japanese Psychological Research*, 51, 197–208.
766

767 Kondo, N., Izawa, E., & Watanabe, S. (2010). Perceptual mechanism for vocal
768 individual recognition in jungle crows (*Corvus macrorhynchos*): contact call
769 signature and discrimination. *Behaviour*, 147, 1051–1072.
770 doi:10.1163/000579510X505427

771

772 Kreiman, J. & Sidtis, D. (2011). *Foundations of Voice Studies: An*
773 *Interdisciplinary Approach to Voice Production and Perception*. Blackwell
774 Publishing, Oxford, UK.

775

776 Lavan, N., Burton, A.M., Scott, S.K. & McGettigan, C. (2019). Flexible voices:
777 identity perception from variable vocal signals. *Psychonomic Bulletin &*
778 *Review*, 26, 90–102. doi:10.3758/s13423-018-1497-7

779

780 Leliveld, L.M.C., Scheumann, M. & Zimmermann, E. (2011). Acoustic
781 correlates of individuality in the vocal repertoire of a nocturnal primate
782 (*Microcebus murinus*). *The Journal of the Acoustical Society of America*, 129,
783 2278-2288. doi:10.1121/1.3559680

784

785 Lemasson, A. & Hausberger, M. (2011). Acoustic variability and social
786 significance of calls in female Campbell's monkeys (*Cercopithecus campbelli*
787 *campbelli*). *The Journal of the Acoustical Society of America*, 129, 3341-3352.
788 doi:10.1121/1.3569704s

789

790 Levréro, F. & Mathevon, N. (2013). Vocal signature in wild infant chimpanzees.
791 *American Journal of Primatology*, 75, 324-332. doi:10.1002/ajp.22108

792

793 Manser, M. B., Jansen, D. A., Graw, B., Hollén, L. I., Bousquet, C. A., Furrer,
794 R. D., & Le Roux, A. (2014). Vocal complexity in meerkats and other
795 mongoose species. In: *Advances in the Study of Behavior, Vol. 46* (pp. 281-
796 310). Academic Press.

797

798 Marler, P. (1977). The structure of animal communication sounds. In: T.H.
799 Bullock (Ed.), *Recognition of complex acoustic signals*. (pp. 17–35). Berlin:
800 Springer.

801

802 Mathevon, N., Koralek, A., Weldele, M., Glickman, S.E. & Theunissen, F.E.
803 (2010). What the hyena's laugh tells us: sex, age dominance and individual
804 signature in the giggling call of *Crocuta crocuta*. *BMC Ecology*, 10.
805 doi:10.1186/1472-6785-10-9

806

807 Matrosova, V.A., Blumstein, D.T., Volodin, I.A. & Volodina, E.V. (2011). The
808 potential to encode sex, age, and individual identity in the alarm calls of three
809 species of Marmotinae. *Naturwissenschaften*, 98, 181–192.
810 doi:10.1007/s00114-010-0757-9

811

812 Maynard-Smith, J. & Harper, D. (2003). *Animal Signals*. New York: Oxford
813 University Press.

814

815 Mitani, J. C., & Nishida, T. (1993). Contexts and social correlates of long-
816 distance calling by male chimpanzees. *Animal Behaviour*, 45(4), 735-746.
817

818 Mitani, J. C., Gros-Louis, J., & Macedonia, J. M. (1996). Selection for acoustic
819 individuality within the vocal repertoire of wild chimpanzees. *International*
820 *Journal of Primatology*, 17, 569–583. doi:10.1007/BF02735192
821

822 Mitani, J. C., & Stuht, J. (1998). The evolution of nonhuman primate loud calls
823 : acoustic adaptation for long-distance transmission. *Primates*, 39, 171–182.
824

825 Morton, E.S. (1977). On the occurrence and significance of motivation-
826 structural rules in some bird and mammal sounds. *American Naturalist*, 111,
827 855–869.
828

829 Mundry R. & Sommer C. (2007). Discriminant function analysis with
830 nonindependent data: consequences and an alternative. *Animal Behaviour*,
831 74, 965-976. doi:10.1016/j.anbehav.2006.12.028
832

833 Pisanski, K., Nowak, J., & Sorokowski, P. (2016a). Individual differences in
834 cortisol stress response predict increases in voice pitch during exam stress.
835 *Physiology & Behavior*, 163, 234e238. doi: 10.1016/j.physbeh.2016.05.018
836

837 Pisanski, K., Cartei, V., McGettigan, C., Raine, J., & Reby, D. (2016b). Voice
838 Modulation: A Window into the Origins of Human Vocal Control? *Trends in*
839 *Cognitive Sciences*, 20, 304-318. doi:10.1016/j.tics.2016.01.002
840

841 Pisanski, K., Raine, J., & Reby, D. (2020). Individual differences in human
842 voice pitch are preserved from speech to screams, roars and pain cries. *Royal*
843 *Society open science*, 7(2), 191642.
844

845 Price, T., Arnold, K., Zuberbuhler, K. & Semple, S. (2009). Pyow but not hack
846 calls of the male putty-nosed monkey (*Cercopithecus nictitans*) convey
847 information about caller identity. *Behaviour*, 146, 871-888.
848 doi:10.1163/156853908X396610
849

850 Proops, L., McComb, K. & Reby, D. (2012). Cross-modal individual recognition
851 in domestic horses (*Equus caballus*). *Proceedings of the National Academy of*
852 *Sciences*, 106, 947-951. doi: 10.1073/pnas.0809127105
853

854 Raine, J., Pisanski, K., & Reby, D. (2017). Tennis grunts communicate
855 acoustic cues to sex and contest outcome. *Animal Behaviour*, 130, 47-55.
856

857 Rendall, D., Rodman, P. S. & Emond, R. E. (1996). Vocal recognition of
858 individuals and kin in free-ranging rhesus monkeys. *Animal Behaviour*, 51,
859 1007-1015.

860

861 Rendall, D., Owren, M. J., & Rodman, P. S. (1998). The role of vocal tract
862 filtering in identity cueing in rhesus monkey (*Macaca mulatta*) vocalisations.
863 *Journal of Acoustic Society of American*, *103*(1), 602–614.

864

865 Robbins, M. M., Bermejo, M., Cipolletta, C., Magliocca, F., Parnell, R.J. &
866 Stokes, E. (2004). Social structure and life-history patterns in western gorillas
867 (*Gorilla gorilla gorilla*), *American Journal of Primatology*, *64*, 145-159.
868 doi:10.1002/ajp.20069

869

870 Robisson, P., Aubin, T. & Bremond, J. C. (1993). Individuality in the voice of
871 the emperor penguin, *Aptenodytes forsteri*: Adaptation to a noisy environment.
872 *Ethology*, *94*, 279-290. doi:10.1111/j.1439-0310.1993.tb00445.x

873

874 Rubow, J., Cherry, M. I., & Sharpe, L. L. (2017). A comparison of individual
875 distinctiveness in three vocalisations of the dwarf mongoose (*Helogale*
876 *parvula*). *Ethology*, *124*, 45–53. <https://doi.org/10.1111/eth.12698>

877

878 Salmi, R., Hammerschmidt, K. & Doran-Sheehy, D.M. (2014). Individual
879 distinctiveness in call types of wild western female gorillas. *PLOS One*, *9*,
880 e101940. doi:10.1371/journal.pone.0101940

881

882 Schamberg, I., Cheney, D. L., Clay, Z., Hohmann, G., & Seyfarth, R. M.
883 (2016). Call combinations, vocal exchanges and interparty movement in wild
884 bonobos. *Animal Behaviour*, *122*, 109-116.
885

886 Schamberg, I., Cheney, D. L., Clay, Z., Hohmann, G., & Seyfarth, R. M.
887 (2017). Bonobos use call combinations to facilitate inter-party travel
888 recruitment. *Behavioral ecology and sociobiology*, *71*(4), 75.
889

890 Searcy, W. A., Anderson, R. C., & Nowicki, S. (2006). Bird song as a signal of
891 aggressive intent. *Behavioral Ecology and Sociobiology*, *60*, 234-241.
892

893 Silk, J. B. (2002). Grunts, girneys, and good intentions: The origins of strategic
894 commitment in nonhuman primates. *Evolution and the capacity for*
895 *commitment*, *3*, 138-157.
896

897 Snowdon, C.T & M. Hausberger (1997). *Social influences on vocal*
898 *development*. Cambridge, U.K.: Cambridge University Press.
899

900 Soltis, J., Leong, K., & Savage, A. (2005). African elephant vocal
901 communication II: Rumble variation reflects the individual identity and
902 emotional state of callers. *Animal Behaviour*, *70*, 589–599.
903 doi:10.1016/j.anbehav.2004.11.016
904

905 Surbeck, M. & Hohmann, G. (2013). Intersexual dominance relationships and
906 the influence of leverage on the outcome of conflicts in wild bonobos (*Pan*
907 *paniscus*). *Behavioral Ecology and Sociobiology*, *67*, 1767–1780. doi:
908 10.1007/s00265-013-1584-8

909

910 Taylor, A. M., Charlton, B. D., & Reby, D. (2016). Vocal production by
911 terrestrial mammals: Source, filter, and function. In: R. A. Suthers, W. T. Fitch,
912 R. R. Fay, & A. N. Popper (Eds.), *Vertebrate sound production and acoustic*
913 *communication* (pp. 229e259). New York, NY: Springer International
914 Publishing. doi: 10.1007/978-3-319-27721-9_8.

915

916 Tallet, C., Linhart, P., Policht, R., Hammerschmidt, K., Šimeček, P., Kratinova,
917 P., & Špinka, M. (2013). Encoding of situations in the vocal repertoire of piglets
918 (*Sus scrofa*): a comparison of discrete and graded classifications, *PLOS One*,
919 *8*, e71841.

920

921 Vannoni, E. & McElligott, A. G. (2007). Individual acoustic variation in fallow
922 deer (*Dama dama*) common and harsh groans: A source-filter theory
923 perspective. *Ethology*, *113*, 223–234. doi:10.1111/j.1439-0310.2006.01323.x

924

925 Vignal, C., Mathevon, N., & Mottin, S. (2004). Audience drives male songbird
926 response to partner's voice. *Nature*, *430*, 448-451.

927

928 Voigt-Heucke, S. L., Taborsky, M. & Dechmann, D.K.N. (2010). A dual function
929 of echolocation: bats use echolocation calls to identify familiar and unfamiliar
930 individuals. *Animal Behaviour*, *80*, 59–67. doi:10.1016/j.anbehav.2010.03.025
931

932 White, F. J. (1996). Comparative socio-ecology of *Pan paniscus*. In W.C.
933 McGrew, L.F. Marchant & T. Nishida (Eds.), *Great Ape Societies* (pp. 29–41).
934 Cambridge, U.K.: Cambridge University Press.
935

936 White, F. J., Waller, M., Boose, K., Merrill, M.Y. & Wood, K.D. (2015). Function
937 of loud calls in wild bonobos. *Journal of Anthropological Sciences*, *93*, 1-13.
938 doi:10.4436/JASS.93003
939

940 Wiley, R. H. (2013). Specificity and multiplicity in the recognition of individuals:
941 implications for the evolution of social behaviour. *Biological Reviews*, *88*, 179–
942 195.
943
944
945

946 Table 1. Automatic and manual acoustic parameters
 947
 948

| Acoustic parameters | Description/Calculation |
|------------------------------------|--|
| F ₀ -Start | Fundamental frequency at beginning of the call (Hz) |
| F ₀ -End | Fundamental frequency at end of the call (Hz) |
| F ₀ -Peak | Highest frequency reached on the fundamental (in Hz) |
| F ₀ -Peak Time | Point over the duration of the call at which F ₀ -Peak is reached. Manually calculated as a proportion: time of F ₀ -Peak(s) / Call Duration (s) |
| Call Duration | Length of call (s) |
| Ascending Slope | Calculated as: $(F_0\text{-Peak} - F_0\text{-Start}) / (F_0\text{-Peak Time} - 0)$ |
| Descending Slope | Calculated as: $(F_0\text{-End} - F_0\text{-Peak}) / (1 - F_0\text{Peak Time})$ |
| Slope- F ₀ Start to Mid | Calculated as: $(F_0 \text{ at midpoint of call duration} - F_0\text{-Start}) / (\text{Time at midpoint of call duration} - 0)$ |
| Slope- F ₀ Mid to End | Calculated as: $(F_0\text{-End} - F_0 \text{ at midpoint of call duration}) / (\text{Call duration} - \text{Time at midpoint of call duration})$ |
| Q25 Frequency | The frequency at which the call is divided into two intervals, the first quartile contains 25% of the call's energy (Hz) |
| Q25 Time | The time at which the call is divided into two intervals, the first quartile contains 25% of the call's energy (s) |
| Q50 Frequency | The median - the frequency at which the call is |

| | |
|--------------------------|---|
| | divided into two frequency intervals of equal energy (Hz) |
| Q75 Frequency | The third quartile – contains 75% of the call's energy (Hz) |
| Q75 Time | The third quartile – contains 75% of the call's energy (s) |
| Maximum Frequency | The frequency at which the maximum energy occurs in the call (Hz) |
| Maximum Time | The first time point along the call where maximum amplitude occurs on waveform (s) |

949

950 Non-bold text are the manually measured or calculated measurements, and

951 bold text are automatically calculated by the Raven program.

952

953 Table 2. Call type usage in each context
 954
 955

| Percentage of calls given in each context | | | | | |
|---|--------------|--------------|--------------|--------------|--------------|
| Contexts | High-hoot | Bark | Soft Bark | Peep-yelp | Peep |
| Aggression | 23.12 | 19.03 | 2.66 | 0.00 | 0.00 |
| Change in environment | 18.02 | 13.46 | 9.93 | 2.86 | 1.58 |
| External Event (Alarm) | 13.81 | 6.96 | 2.18 | 2.14 | 0.79 |
| Pre-feeding | 38.74 | 34.34 | 21.55 | 17.38 | 21.74 |
| Feeding | 1.80 | 18.10 | 38.26 | 36.19 | 28.46 |
| Forage | 4.50 | 3.25 | 18.40 | 15.95 | 9.09 |
| Groom | 0.00 | 0.93 | 1.45 | 6.43 | 9.88 |
| Contact-general | 0.00 | 3.94 | 5.57 | 19.05 | 28.46 |

956

957 The two contexts in which each call type is most commonly used are in bold.
 958 (Percentages presented for clarity only; chi-square results were generated by
 959 analysing observed vs. expected values of call rates.)

960 Table 3. Potential for individual coding (PIC)
 961
 962

| PIC values by Call Types | | | | | | |
|----------------------------------|--------------------------|-------------|-------------|-------------|-------------|-------------|
| Individual Parameters | Overall PIC | | | | | |
| | value for each parameter | Peep | Peep -yelp | Soft bark | Bark | High-hoot |
| F ₀ -Start | 1.20 | 1.23 | 1.24 | 1.12 | 1.29 | 1.38 |
| F ₀ -End | 1.11 | 1.20 | 1.18 | 1.20 | 1.19 | 1.43 |
| F ₀ -Peak | 1.20 | 1.21 | 1.25 | 1.12 | 1.25 | 1.46 |
| F ₀ -Peak Time | 1.08 | 0.91 | 1.03 | 1.04 | 1.06 | 1.18 |
| Call Duration | 1.16 | 1.05 | 1.13 | 1.20 | 1.19 | 1.28 |
| Ascending Slope | 1.22 | 0.97 | 1.15 | 1.06 | 1.19 | 1.14 |
| Descending Slope | 1.03 | 1.06 | 1.17 | 1.08 | 1.02 | 1.21 |
| Q25 Frequency | 1.17 | 1.17 | 1.11 | 1.12 | 1.17 | 1.38 |
| Q25 Time | 1.07 | 0.99 | 1.01 | 1.03 | 1.15 | 1.06 |
| Q50 Frequency | 1.12 | 1.15 | 1.11 | 1.10 | 1.22 | 1.36 |
| Q75 Frequency | 1.14 | 1.17 | 1.08 | 1.18 | 1.23 | 1.34 |
| Q75 Time | 1.09 | 1.00 | 1.05 | 1.06 | 1.12 | 1.14 |
| Maximum Frequency | 1.08 | 1.12 | 1.08 | 1.08 | 1.11 | 1.28 |
| Maximum Time | 1.10 | 1.02 | 1.03 | 1.03 | 1.10 | 1.08 |
| Overall PIC for call type | | 1.03 | 1.10 | 1.08 | 1.18 | 1.22 |

963

964 Investigation of the potential for individuality in each call type as well as each
965 parameter across all call types and within each call type separately. A PIC
966 greater than or equal to 1 suggests potential for individuality. For each call
967 type, the parameter with the highest PIC is in bold.
968

969 Figure 1. Acoustic and temporal measurements. Example of manual
970 measurements taken on a soft bark call: a= F0-Start, b=F0-Midpoint (not
971 included in analysis but used to calculate other measurements), c=F0-Peak, d=
972 F0-End. Call duration=time at b) – time at a).

973

974

975 Figure 2. Spectrographic illustrations of the five most common call types of the
976 bonobo repertoire. Spectrogram images depict: 1) High-hoot; 2) Bark; 3) Soft
977 bark; 4) Peep-yelp; 5) Peep. Panel 6) shows the confusion matrix obtained
978 from the permuted DFA classifying the five call types, which are labelled as
979 just listed. The classification rate was of 60% for High-hoot, 44% for Bark, 48%
980 for Soft bark, 58% for Peep-yelp and 62% for Peep. On the confusion matrix,
981 the diagonal shows the rate at which a call type was correctly assigned –
982 specifically, what percentage the actual call type and the predicted call type
983 match. The brightness of each diagonal yellow square corresponds to the
984 strength of classification. The off-diagonal cells show percentage of
985 misclassification for each call type and which other call type they were
986 misclassified as.

987

988

989 Figure 3. Acoustic gradation and individual signature coding in the bonobo
990 vocal repertoire. Top: The scatterplot illustrates the graded nature of the
991 acoustic properties of the five call types. The Table displays the factor loadings

992 on Discriminant function 1 and 2. White asterisks mark the centroid of each
993 call type. Bottom graph: The mean percentage of correct classification for
994 individual identity is reported for each call type, illustrating that the reliability of
995 the vocal signature increases from peeps to high-hoots.

996

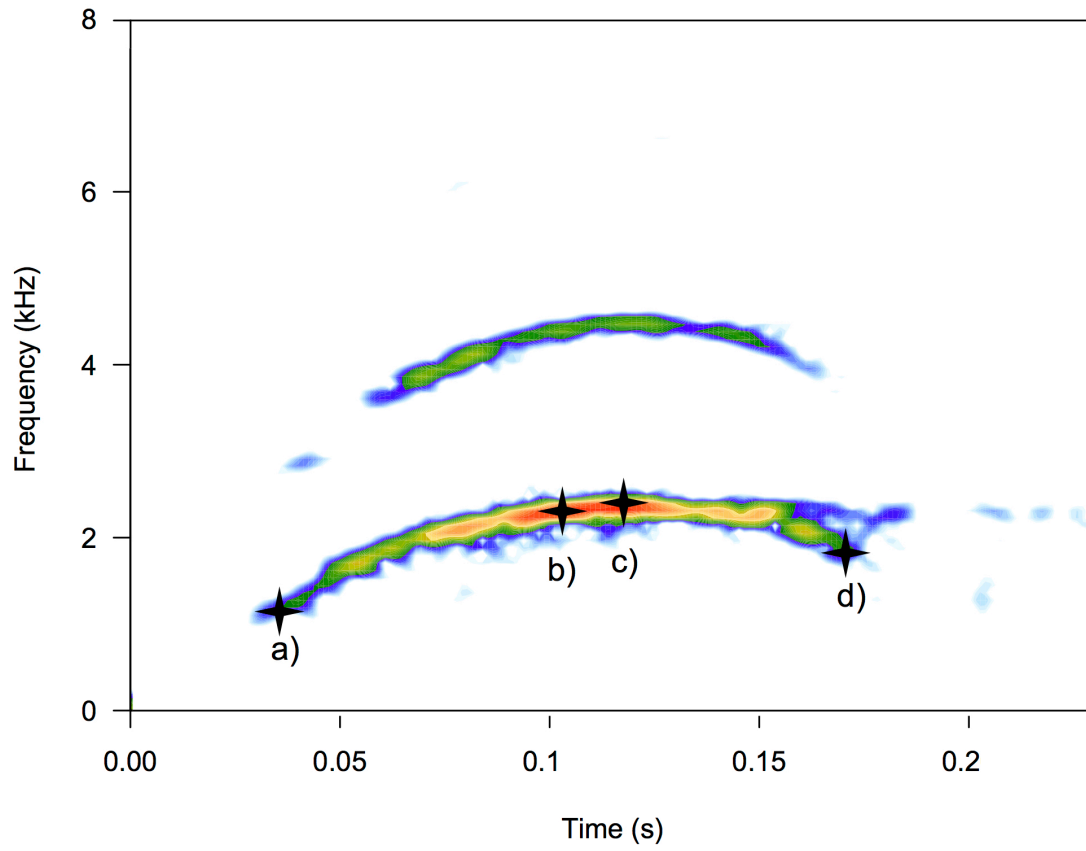
997

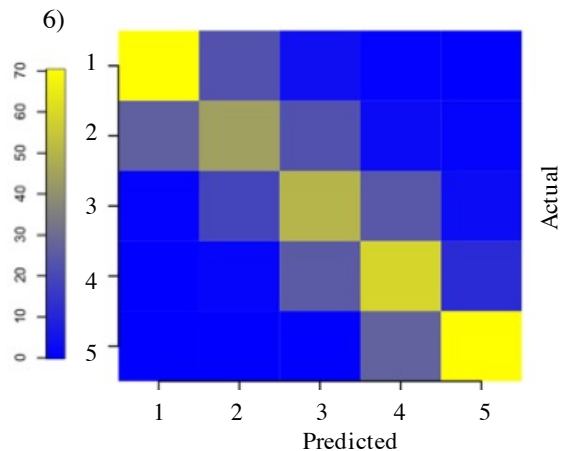
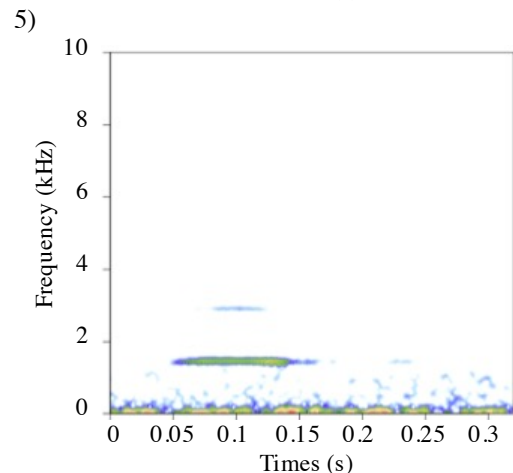
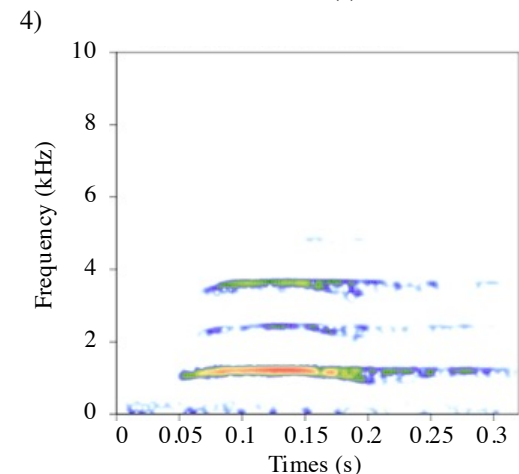
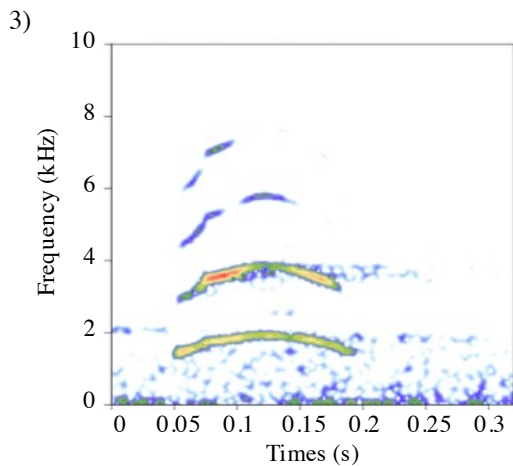
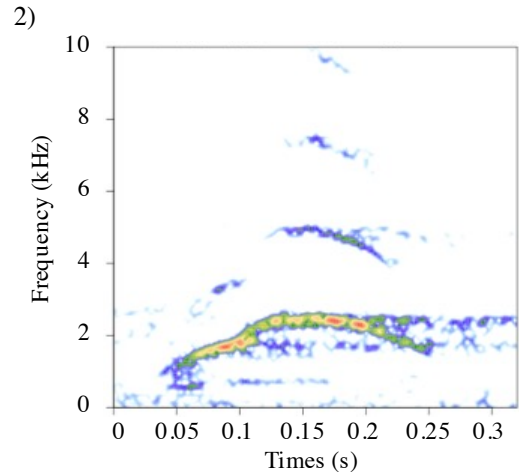
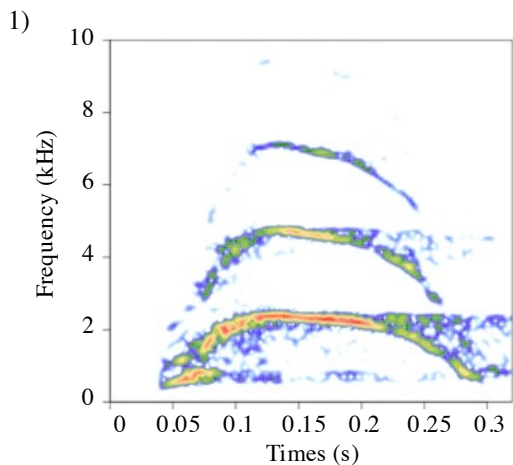
998 Figure 4. Individual vocal signatures in five different call types. 1) High-hoots
999 (mean correction classification rate: 53%); 2) Barks (mean correction
1000 classification rate: 44%; 3) Soft Barks (mean correction classification rate:
1001 30%); 4) Peep-yelps (mean correction classification rate: 30%); 5) Peeps
1002 (mean correction classification rate: 23%). Each confusion matrix shows the
1003 results of five separate permuted DFAs investigating the strength individuality
1004 in each call type. The confusion matrix shows the probability that an
1005 individual's calls were correctly classified after 100 iterations – the legend
1006 shows the percent accuracy for each individual (the accurate percentages are
1007 given in Supplementary Table A4; Random classification at 10%). Details on
1008 each individual, age, sex, rank, zoo and number call contributed to each
1009 analysis can be found in Supplemental Table A1.)

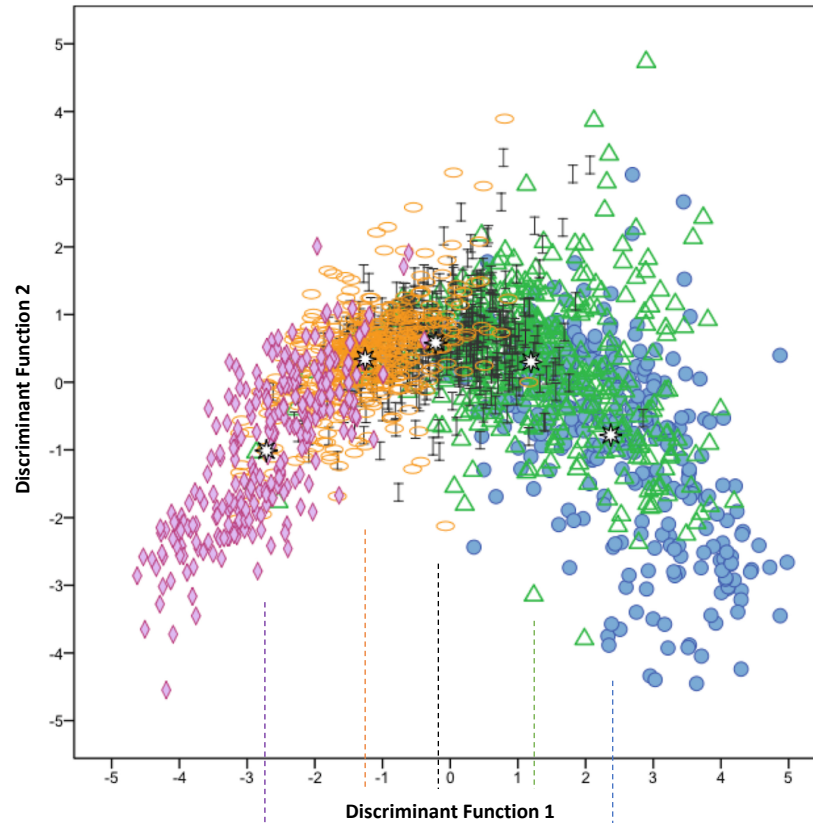
1010

1011

1







Call Type

- High-hoot
- △ Bark
- I Soft Bark
- Peep-yelp
- ◇ Peep
- ★ Group Centroid

| | Function 1 | Function 2 |
|--------------------------------|---------------|--------------|
| Fo-Start | -0.121 | 0.436 |
| Fo-End | -0.015 | 0.146 |
| Fo-Maximum | 0.616 | 0.024 |
| Fo-Maximum proportion | 0.076 | 0.293 |
| Call Duration | 0.445 | -0.003 |
| <u>Ascending Slope</u> | 0.59 | -0.452 |
| <u>Descending Slope</u> | -0.612 | 0.151 |
| <u>Slope-Start to Midpoint</u> | 0.543 | -0.183 |
| <u>Slope-Midpoint to End</u> | -0.467 | -0.166 |
| <u>Q1 Frequency</u> | 0.244 | 0.111 |
| <u>Q1 Time</u> | 0.195 | 0.068 |
| <u>Q3 Frequency</u> | 0.299 | -0.012 |
| <u>Q3 Time</u> | 0.34 | -0.044 |
| <u>Centre Frequency</u> | 0.303 | 0.006 |
| <u>Maximum Frequency</u> | 0.214 | 0.057 |
| <u>Maximum Time</u> | 0.238 | 0.019 |

