

Manuscript version: Author's Accepted Manuscript

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

Persistent WRAP URL:

<http://wrap.warwick.ac.uk/115163>

How to cite:

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

Copyright and reuse:

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

Copyright © and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable the material made available in WRAP has been checked for eligibility before being made available.

Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge. Provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Publisher's statement:

Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk.

1 **Article Title:**

2 Combinatoriality in the vocal systems of non-human animals

3

4 **Article Type:**

5 Overview

6

7 **Authors:**

First author

Sabrina Engesser *

ORCID iD: 0000-0003-0425-2179

Email: sabrina.engesser@outlook.com

Department of Comparative Linguistics, University of Zurich, Zurich, Switzerland

No conflict of interest

Second author

Simon William Townsend

ORCID iD: 0000-0003-1504-1801

Email: simonwilliam.townsend@uzh.ch

Department of Comparative Linguistics, University of Zurich, Zurich, Switzerland

Department of Psychology, University of Warwick, Coventry, United Kingdom

No conflict of interest

8

9 **Abstract**

10 A key challenge in the field of human language evolution is the identification of the selective
11 conditions that gave rise to language's generative nature. Comparative data on non-human
12 animals provides a powerful tool to investigate similarities and differences among non-
13 human and human communication systems and to reveal convergent evolutionary
14 mechanisms. In this article, we provide an overview of the current evidence for combinatorial
15 structures found in the vocal system of diverse species. We show considerable structural
16 diversity exists across and within species in the forms of combinatorial structures used. Based
17 on this we suggest that a fine-grained classification and differentiation of combinatoriality is
18 a useful approach permitting systematic comparisons across animals. Specifically, this will
19 help to identify factors that might promote the emergence of combinatoriality and, crucially,
20 whether differences in combinatorial mechanisms might be driven by variations in social and
21 ecological conditions or cognitive capacities.

22

23 **Graphical/Visual Abstract and Caption**

24 Use Fig. 1.

25 The article provides an overview of combinatorial structures characterising non-human vocal
26 systems and how they might relate to human linguistic structures.

27 **Introduction**

28 Language is considered a hallmark of the human species (Darwin, 1871; Hockett, 1960). Our
29 ability to purposefully communicate (*intentionality*) meaningful information (*semanticity*) in
30 an open-ended way (*generativity/productivity*) appears to set us apart from the rest of the
31 animal kingdom (Hockett, 1960). The question of the origin of this capacity has puzzled
32 scholars ever since Darwin drew attention towards the obvious differences between human
33 language and animal communication (Darwin, 1871), and much effort has been put into
34 elucidating language's uniqueness and which language-specific traits might be shared with
35 non-human animals (Hauser, Chomsky, & Fitch, 2002; Pinker & Jackendoff, 2005). As a
36 result, comparative studies conducted over the last decades have revealed that several
37 language-specific components do have analogues in animal systems (Hauser et al., 2002;
38 Pinker & Jackendoff, 2005). For example, both the ability to intentionally communicate
39 information (c.f. vocal flexibility (Marler, Dufty, & Pickert, 1986; Zuberbühler, 2008)) and to
40 assign acoustic labels to external stimuli (c.f. functional referential signals (Seyfarth, Cheney,
41 & Marler, 1980; Suzuki, 2016; Townsend & Manser, 2013)) have been argued to be present
42 in diverse animal species, providing important insights into the evolutionary drivers of these
43 abilities (Christiansen & Kirby, 2003; Hauser et al., 2002) (but see Sievers and Gruber
44 (2016); Townsend et al. (2016); Wheeler and Fischer (2012) for further reading and ongoing
45 debates). However, one critical component remains obscured: the evolutionary origin of
46 language's generative nature (Hauser, 1997; Humboldt, 1999; Nowak, 2000; Zuidema & de
47 Boer, 2018).

48 Language's generativity is the product of its double articulation, which involves the
49 combination of acoustic elements and units on two different levels (Hockett, 1960). On the
50 *phonological layer (combinatorial phonology)*, a limited number of meaningless speech
51 sounds (e.g. /a/ /t/ /k/ /h/) can be (re-)combined to create a theoretically infinite array of

52 morphemes or words, (e.g. “at”, “cat”, “hat”), whereby those sounds that serve to distinguish
53 meaning are classified as phonemes (e.g. /k/ & /h/ when contrasting “cat” & “hat”) (Chomsky
54 & Halle, 1968). At the higher *syntactic layer* (*compositional semantics*) these meaning-
55 encoding components can then be assembled into larger structures, with the structure’s
56 meaning being derived from its individual components and the rules that govern their
57 organisation (e.g. “the cat in the hat”) (Chomsky, 1957; Hurford, 2007, 2012b). Traditionally,
58 the search for comparative examples of phonological and syntactic features outside of human
59 language has i) focused on song-driven systems, arguably the most obvious and complex
60 forms of sound combinations, or ii) searched for homolog examples in primates, under the
61 assumption that our generative capacities are adaptations of pre-existing traits shared with our
62 closest-living relatives (Berwick, Okanoya, Beckers, & Bolhuis, 2011; Hauser et al., 2002;
63 Yip, 2006; Zuberbühler, 2018). Only more recently has this comparative approach been
64 extended to the discrete vocal systems of distantly related species (as opposed to song-driven
65 or primate systems). As a result, the growing body of work suggests that the ability to
66 combine acoustic segments into larger structures is by no means a rare phenomenon in the
67 animal kingdom, with some combinatorial structures providing analogues to linguistic
68 structures in human language and others lacking any apparent resemblance.

69 In line with the accumulating evidence of animal combinatoriality, the objective of
70 this review is to provide an overview of diverse forms of naturally produced vocal
71 combinations found in the communication systems of non-human animals (but see also
72 Zuberbühler (2018) for a synopsis on combinatoriality in primates; as well as ten Cate (2017)
73 and Fitch (2018) for recent reviews on cognitive abilities related to the computation of
74 combinatorial sequences). We aim to illustrate the diverse forms combinatoriality can take in
75 animal vocal systems and the extent to which these might resemble linguistic structures in
76 human language. In doing so we hope to stimulate further research investigating the selective

77 conditions and underlying mechanisms that drive the emergence of generative vocal
78 mechanisms characterising animal vocal systems.

79

80 **Animal combinatoriality**

81 Within and across species, the combinatorial forms characterising animal vocal systems
82 involve an interesting degree of structural and semantic complexity (Hurford, 2012a; Marler
83 & Slabbekoorn, 1999). Classically, vocal sequences produced by animals have been
84 categorised as either a form of *phonological syntax* (or *combinatoriality*) if the sequences are
85 composed of so-called *meaningless* elements; or alternatively as a form of *lexical syntax* (or
86 *compositionality*) if a sequence constitutes a combination of, and derives its overall meaning
87 from, individually meaningful signals (Hurford, 2007; Marler, 1977). While this binary
88 discrimination initially has proven useful in defining crude similarities and differences among
89 animal and human communication systems (Collier, Bickel, van Schaik, Manser, &
90 Townsend, 2014; Hurford, 2007; Marler, 1977), recent research indicates that many animal
91 vocal sequences cannot easily be assigned to one or the other layer. Instead, both animals and
92 human combinatorial systems can be decomposed into finer, transitional forms (e.g.
93 affixation) or sometimes even feature both phonological and syntactic aspects (e.g. music)
94 (Rohrmeier, Zuidema, Wiggins, & Scharff, 2015; Townsend, Engesser, Stoll, Zuberbühler, &
95 Bickel, 2018).

96 In the following sections, we review and categorise current examples of animal vocal
97 combinations, in turn enabling a more precise classification of the existing diversity of animal
98 vocal constructs (see Fig. 1 for graphical illustration). Acoustic segments will be referred to
99 as *meaningless* if they represent distinguishable elements that are not emitted in isolation, and
100 hence are unlikely to transfer functionally relevant information, or *meaningful* if their

101 production is context-specific and elicit predictable responses in receivers suggesting they
102 serve a distinct function (Marler, 2000).

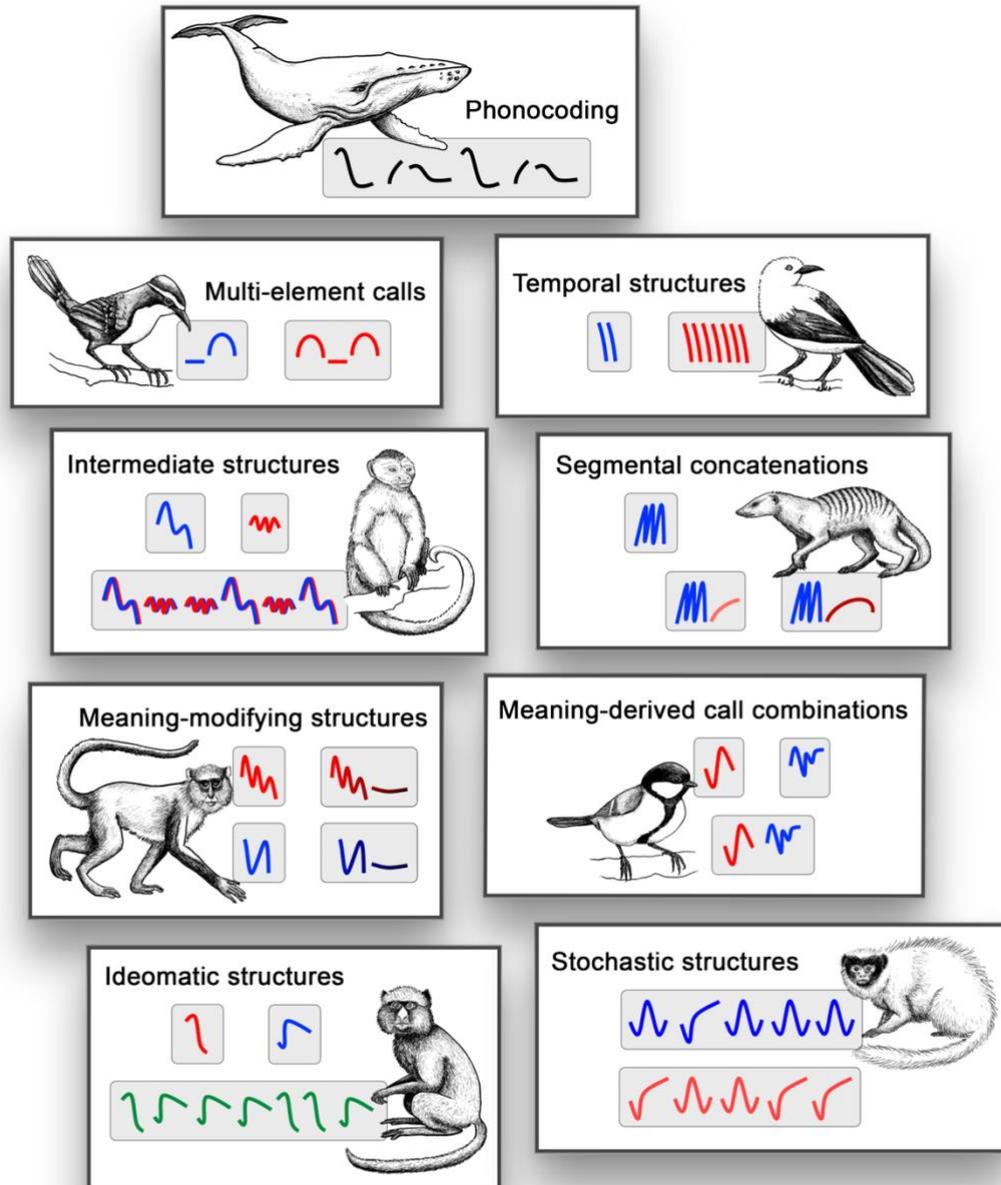


Fig. 1: Graphical illustration of combinatorial structures in non-human vocal systems. Each combinatorial structure shows one representative species. Shapes distinguish sound elements, colours distinguish meaning (black indicates absence of functional- or context-specific meaning). **Phonocoding**: combination of meaningless elements into sequence that lacks functional- or context-specific meaning (e.g. whales, songbirds). **Multi-element calls**: combination and reuse of meaningless elements to generate context-specific/functionally meaningful calls (e.g. chestnut-crowned babblers). **Temporal structures**: meaning-differentiating temporal variation (e.g. number of element repetitions) within a string of repeated sounds (e.g. pied babblers, Mexican free-tailed bats). **Intermediate structures**: combination of meaningful calls into sequence reflecting intermediate stages experienced by the caller (e.g. wedge-capped capuchins, gorillas). **Segmental concatenations**: concatenation of invariable, stand-alone segment (generally identity-encoding) with variable, bound segment (generally motivation/behaviour-encoding) (e.g. banded mongooses, Diana monkeys). **Meaning-modifying structures**: combination of individually meaningful (variable) call with (invariable) meaning-modifying affix (e.g. Campbell's monkeys). **Meaning-derived call combinations**: combination of meaningful calls into sequence with derived meaning (e.g. Japanese tits, pied babblers). **Idiomatic structures**: combination of meaningful calls into sequence with unrelated meaning (e.g. putty-nosed monkeys). **Stochastic structures**: sequence whose meaning is encoded by proportional sound/call contribution at particular parts of the sequence (e.g. black-fronted titi monkeys). Animal drawings by Zinaida Bogdanova.

104 *Phonocoding (meaning-devoid sound combinations)*

105 A wide range of animals, including songbirds, bats, gibbons, hyraxes and whales, combine
106 meaningless sound elements into higher-order, often hierarchically structured, sequences or
107 songs (Bohn, Smarsh, & Smotherman, 2013; Catchpole & Slater, 1995; Geissmann, 2002;
108 Kershenbaum, Ilany, Blaustein, & Geffen, 2012; Payne & McVay, 1971). Such combinatorial
109 sequences predominantly function in territorial or courtship display (Catchpole & Slater,
110 1995; Marler & Slabbekoom, 1999; Mitani & Marler, 1989), to facilitate recognition between
111 individuals or groups (Antunes et al., 2011; Briefer, Rybak, & Aubin, 2013; Holland,
112 Dabelsteen, & Paris, 2000; Nousek, Slater, Wang, & Miller, 2006; Schulz, Whitehead, Gero,
113 & Rendell, 2011), or to strengthen the bonding among partners or groups (Janik & Slater,
114 1997; King & Janik, 2013). Although, songs have traditionally been defined as phonological
115 syntax (Marler, 1977), a more detailed analysis would suggest they neither qualify as
116 phonology nor as syntax. Firstly, albeit composed of meaningless elements, songs lack a
117 functionally- or context-specific meaning (i.e. songs lack propositional semantics), but
118 instead serve to more broadly signal caller attributes (independent of context and content),
119 hence contrasting with human phonology (Berwick et al., 2011; Rendall, 2013). Secondly,
120 although on the surface level, a song's internal structuring (notes being arranged into
121 syllables, motifs, phrases etc.) resembles language's syntactic architecture, songs lack the
122 combinatorial semantics of language's syntactic layer with the precise arrangement of sounds
123 being irrelevant for a song's informational content (Rendall, 2013; Rohrmeier et al., 2015).
124 Nevertheless, intriguing parallels can be drawn between animal songs and human music, both
125 being rich in structural complexity, yet, with combinatorial variation being unimportant for
126 meaning-differentiation (Bowling & Fitch, 2015; Rohrmeier et al., 2015). As such, animal
127 songs are suggested to offer a model system to study the evolutionary origins of human music

128 and its structural complexity, as well as music’s developmental and neural basis (Berwick et
129 al., 2011; Rohrmeier et al., 2015).

130

131 *Multi-element calls (meaningful sound combinations)*

132 In contrast to the meaning-*devoid* sound combinations characterising songs, some species
133 have been described to also combine meaningless sound elements to generate *functionally*
134 *relevant* vocalisations. Among the best known; chickadees, tits and titmice (Paridae family)
135 produce multi-element “chick-a-dee” or “chicka” calls, with supposedly meaningless
136 elements being arranged in a stereotyped order, and with omissions or duplications of
137 individual sounds generating dozens of call variants (Ficken, Hailman, & Hailman, 1994;
138 Hailman, 1989; Hailman, Ficken, & Ficken, 1985; Suzuki, 2013). Although there seems to be
139 a degree of context specificity with some call variants being more likely produced during
140 certain behaviours, potentially linked to locomotion and movement (Ficken et al., 1994;
141 Hailman & Ficken, 1986; Hailman et al., 1985; Suzuki, 2013) or eliciting different responses
142 in receivers (Clucas, Freeberg, & Lucas, 2004; Freeberg & Lucas, 2002), the extent to which
143 the variants encode qualitatively different information remains to be tested (Ficken et al.,
144 1994; Freeberg & Lucas, 2012; Hailman et al., 1985). Outside the Paridae family, chestnut-
145 crowned babblers (*Pomatostomus ruficeps*) produce a pair of functionally distinct
146 vocalisations composed of two meaningless sounds: AB-flight calls that function to
147 coordinate group movement, and BAB-prompt/provisioning calls that serve to stimulate
148 nestling begging. Specifically, both calls are composed of two perceptibly distinct,
149 meaningless sounds (A and B), with the meaning-differentiation among the two calls being
150 the result of a modification at one position of the calls (i.e. BAB) (Engesser, Crane, Savage,
151 Russell, & Townsend, 2015). While this bears analogies with the phonemic-structuring of
152 words in human language (smallest contrasting elements distinguish semantic meaning; Yip

153 (2006)), of perhaps greater relevance for studies on the combinatorial power in animal
154 communication systems is the shared use of meaningless elements to generate qualitatively
155 distinct signals. Future work should address whether the meaningful vocalisations of other
156 species can be similarly decomposed into smaller, shared elements. It is worth noting that,
157 although vocalisations composed of acoustically isolated elements offer an easy tractable
158 system, calls composed of uninterrupted, meaningless acoustic streams deserve equal
159 consideration.

160

161 *Temporal structures*

162 Besides the combination of *different* sounds, animals further encode information by varying
163 the temporal arrangement of the *same* repeated sound element within a sequence. The most
164 commonly described functions of such temporal modifications is to transfer information on
165 an individual's arousal level experienced during aggressive or predatory encounters
166 (Blumstein, 2007; Courter & Ritchison, 2010; DuBois, Nowicki, & Searcy, 2009; Lemasson,
167 Ouattara, Bouchet, & Zuberbühler, 2010; Mahurin & Freeberg, 2008; Manser, 2001; Rek &
168 Osiejuk, 2012; Sieving, Hetrick, & Avery, 2010; Templeton, Greene, & Davis, 2005;
169 Wheatcroft, 2015). Therefore, information is generally encoded through gradual changes in
170 the number or the rate of repeated elements or changes in the inter-element intervals
171 (Blumstein, 2007; Courter & Ritchison, 2010; DuBois et al., 2009; Lemasson et al., 2010;
172 Mahurin & Freeberg, 2008; Manser, 2001; Rek & Osiejuk, 2012; Sieving et al., 2010;
173 Templeton et al., 2005; Wheatcroft, 2015). Recent work has also demonstrated that temporal
174 changes cannot only encode *quantitative changes* in arousal, but also more qualitative
175 information. For example, work on the alarm call system of colobus monkeys (*Colobus*
176 *guereza* & *C. polykomos*), as well as on the social calls of Mexican free-tailed bats (*Tadarida*
177 *brasiliensis*) and pied babblers (*Turdoides bicolor*) has shown temporal structures can further

178 encode *categorical information* (Bohn, Schmidt-French, Ma, & Pollak, 2008; Engesser,
179 Ridley, & Townsend, 2017; Schel, Candiotti, & Zuberbühler, 2010; Schel, Tranquilli, &
180 Zuberbühler, 2009). Colobus monkeys cluster repeated “roar”-sounds into bouts of different
181 lengths, with bouts then being emitted at different time intervals. Long bouts emitted at larger
182 time intervals are indicative of eagle presence, while short bouts - each introduced by another
183 “snort”-sound - emitted at shorter time intervals are associated with leopard presence (Schel
184 et al., 2010; Schel et al., 2009). However, since the “snort”-element may equally serve in
185 distinguishing the alarm sequences it is unclear to what extent the temporal arrangement is
186 relevant for differentiating the structures’ meaning. Mexican free-tailed bats and pied
187 babblers, on the other hand, produce two qualitatively different vocalisations which are solely
188 discriminated based on the number of element repetitions composing each call. Specifically,
189 Mexican free-tailed bats emit *mono*-syllabic “click”-calls when investigating novel stimuli,
190 and *multi*-syllabic “click”-calls when interacting with conspecifics, with the two call variants
191 potentially serving different functions (Bohn et al., 2008). Similarly, pied babblers produce
192 “cluck”-calls composed of 2-3 repetitions of a short broadband sound, and “purr”-calls
193 composed of 17 ± 10 repetitions of the same sounds. While “clucks” appear to induce a
194 collective group movement based on a decision-making process, “purrs” serve to attract
195 offspring to food sources (Engesser et al., 2017; Radford & Ridley, 2006). Overall, given
196 modifications of a call’s temporal characteristics are assumed to be less constrained than
197 changes of frequency-related features (Janik & Slater, 1997), it is surprising that temporal
198 modifications encoding discrete, categorical information have so far only been demonstrated
199 in few species. One reason may be because such structures do not resemble language’s
200 combinatorial layers, and thus temporal aspects have generally been neglected. Nevertheless,
201 we propose that their simplicity might actually make it a valuable and widespread mechanism
202 applied by animals to encode diverse information.

203

204 *Intermediate/readout call structures*

205 Potentially more widespread, but so far also described in only a handful of species including
206 wedge-capped capuchins (*Cebus olivaceus*), mountain and western gorillas (*Gorilla beringei*
207 *beringei*, *G. gorilla*), and potentially chimpanzees (*Pan schweinfurthii*) and meerkats
208 (*Suricata Suricata*), are combinations of calls that might be interpreted as “online readouts”
209 of the caller’s current motivational state (Collier, Townsend, & Manser, 2017; Crockford &
210 Boesch, 2005; Fedurek, Zuberbühler, & Dahl, 2016; Hedwig, Mundry, Robbins, & Boesch,
211 2014; Robinson, 1984; Zuberbühler, 2018). While, in these instances individual call types can
212 be associated with certain motivational states, combinations appear to reflect intermediate or
213 conflicting interests experienced by the caller during production (e.g. submission vs.
214 aggression) (Cheney & Seyfarth, 1990). Although this does not require signallers to
215 deliberately produce combined structures for informational purposes, receivers may still be
216 able to relate the call’s individual messages and to retrieve potentially relevant information
217 on the conflicting circumstances the caller faces.

218

219 *Segmental concatenations*

220 A few mammalian species concatenate acoustic segments in a seemingly systematic way. For
221 example, Diana monkeys (*Cercopithecus diana*), Campbell’s monkeys (*Cercopithecus*
222 *campbelli campbelli*), banded mongooses (*Mungos mungo*) and dingoes (*Canis lupus dingo*)
223 produce identity-encoding segments which can be given in isolation or concatenated with
224 other distinct or graded elements that correlate with the animal’s motivational/emotional state
225 (e.g. socio-positive/negative context) or its behaviour (e.g. foraging – moving - running)
226 (Candiotti, Zuberbühler, & Lemasson, 2012; Coye, Ouattara, Arlet, Lemmasson, &
227 Zuberbühler, 2018; Coye, Zuberbühler, & Lemasson, 2016; Deaux, Allen, Clarke, &

228 Charrier, 2016; Jansen, Cant, & Manser, 2012). From a linguistic perspective, such segments
229 may resemble morphemes (smallest meaningful units), with the individually distinct elements
230 representing free morphemes that can be produced as a standalone segment, or be conjoined
231 with the behaviour- or motivation-coding (bound) segment (Collier et al., 2014).

232

233 *Meaning-modifying segmental structures (affixation)*

234 Other than the combination of identity- and context-encoding segments, contextually
235 meaningful vocalisations can also be concatenated with meaning-*modifying* acoustic
236 segments (or affixes). Both pied babblers and Campbell's monkeys affix individually
237 meaningful signals with acoustic segments. Although these acoustic segments are never
238 produced in isolation (i.e. are *individually* meaningless), once combined they modify the
239 signal's meaning in a quantitative way and have therefore been argued to carry more of an
240 abstract (intensity-modifying) meaning (Engesser, Ridley, Manser, Manser, & Townsend,
241 2018; Ouattara, Lemasson, & Zuberbühler, 2009b). More precisely, pied babblers produce
242 two variants of a longer sequence that is either composed of repetitions of "A"- or "AB"-note
243 elements. While the single-note structure functions to induce a recruitment of receivers over a
244 short distance to the callers broadcast location, the recruitment-request is intensified through
245 the suffixation of A-notes with B-notes, with the double-note structures inducing a follow of
246 the caller over longer distances (Engesser et al., 2018). However, unlike in human language
247 where affixes modify a signal's meaning in a *predictable* way, the unproductive use of the
248 modifying segment in pied babblers (B is only ever produced in combination with A, but no
249 other sound), renders interpretations regarding its semantic content problematic. In contrast,
250 Campbell's monkeys appear to productively combine a meaning-modifying "-oo" segment
251 with *two* predator specific alarm calls. While in isolation the calls encode leopard or eagle
252 presence, their affixed variants encode general disturbances or unspecific aerial threats

253 respectively (Coye, Ouattara, Zuberbühler, & Lemasson, 2015; Ouattara et al., 2009b).
254 Accordingly, the “-oo” affix appears to modify the predator-specific alarm calls’ meaning in
255 a systematic way (i.e. the affix’s productive use facilitates predictability), by broadening the
256 calls’ meaning (Schlenker et al., 2014; Schlenker, Chemla, & Zuberbühler, 2016). The
257 Campbell’s monkey structure therefore qualifies as a form of basic compositionality, with the
258 modifying segment carrying an abstract meaning, and the combined structure’s meaning
259 reflecting the meaning of its individual parts (Collier et al., 2014; Townsend et al., 2018).

260

261 *Meaning-derived call combinations (semantic compositionality)*

262 Cases where animals combine *stand-alone* meaningful vocalisations into a higher-order
263 structure, whose overall meaning reflects the meaning of its individual parts, are currently
264 only found in birds: specifically, Japanese tits (*Parus minor*) and pied babblers. Both produce
265 alert and recruitment calls, which are combined when mobbing predators (Engesser, Ridley,
266 & Townsend, 2016; Suzuki, Wheatcroft, & Griesser, 2016). In Japanese tits, alert calls elicit
267 vigilance- and recruitment calls approach-behaviour in conspecifics, with the combination
268 eliciting a mixture of both behaviours (Suzuki et al., 2016). Similarly, pied babblers give alert
269 calls to low-urgency threats, and recruitment calls when recruiting group members during
270 collective group travels. Both calls are combined when recruiting group members to mob a
271 terrestrial predator (Engesser et al., 2016). Critically, in both cases the meaning of the
272 combination can be deduced from the meanings of its (individual meaningful) parts, therefore
273 presenting a rudimentary, two-call, compositional structure (Townsend et al., 2018).

274

275 *Idiomatic structures (semantic combinatoriality)*

276 Some monkeys further assemble discrete acoustic units into larger sequences to encode
277 information that appear to be unrelated to the components’ meaning. A textbook example is

278 putty-nosed monkeys (*Cercopithecus nictitans*), which combine two meaningful alarm calls
279 associated with eagle presences and general disturbances, respectively, into longer sequences
280 that elicit group movement in non-predatory contexts (Arnold & Zuberbühler, 2006, 2008,
281 2012). Notably, the resultant structure does not appear to be produced and processed in a
282 compositional way, since the meaning of the whole cannot be derived from its compounds.
283 As such the sequence is suggested to constitute a semantically *combinatorial* or *idiomatic*
284 structure (Arnold & Zuberbühler, 2012; Hurford, 2007, 2012a).

285

286 *Stochastic/proportional structures*

287 Lastly, black-fronted titi monkeys (*Callicebus nigrifrons*) produce sequences composed of
288 two main sound types, with resultant sequences conveying information about the type (raptor
289 vs. carnivore) and location (canopy vs. ground) of a predator (Cäsar, Byrne, Young, &
290 Zuberbühler, 2012; Casär, Zuberbühler, Young, & Byrne, 2013). One analysis is that each of
291 the two sounds might potentially encode a particular meaning, with sequences then
292 representing simple readouts of the current environmental circumstances experienced by a
293 signaller (c.f. *intermediate/readout call sequences*) (Schlenker, Chemla, Cäsar, Ryder, &
294 Zuberbühler, 2016). Alternatively, the overall meaning of a sequence might be derived from
295 the proportional contribution of the individual sounds at particular parts of the sequence
296 (Zuberbühler, 2018). Accordingly, the individual components might lack meaning, but once
297 combined in a rule-governed (i.e. proportional) way, generate meaning (Cäsar et al., 2012;
298 Casär et al., 2013). Similar mechanisms might underlie bonobo (*Pan paniscus*) food call
299 sequences, with sequences potentially encoding the type and quality of food items (Clay &
300 Zuberbühler, 2009, 2011). Though, in both cases, further work needs to investigate whether
301 stable proportions of call contributions can predict context and vice versa.

302

303 More ambiguous and less clearly definable vocal sequences that neither seem to fall in line
304 with a compositional, idiomatic or proportional analysis are produced by gibbons and
305 Campbell’s monkeys. For example, white-handed gibbons (*Hylobates lar*) build structurally
306 distinct sequences specific to social and predatory context from the same acoustic
307 units (Clarke, Reichard, & Zuberbühler, 2006). However, whether the individual units encode
308 meaning, and exactly how the information is derived from the sequences’ (potentially
309 “stochastic”; Zuberbühler (2018)) overall structure remains unclear. Campbell’s monkeys, on
310 the other hand, generate sequences from recombinations of individually meaningful calls and
311 acoustic segments that are only found as part of the larger sequences (Ouattara, Lemasson, &
312 Zuberbühler, 2009a). While the sequences are produced in various contexts, information
313 appears to be encoded through the transition probabilities between, and co-occurrences
314 among, the constituent, permuted parts (Zuberbühler, 2018) (see also Schlenker et al.
315 (2014) for a formal semantic analyses of Campbell’s monkey vocal sequences and alternative
316 interpretations).

317

318 **Conclusion**

319 A major question in the field of language evolution is the origin of language’s generative
320 nature. Tracing its origin, however, poses major difficulties due language’s cumulative
321 evolution not leaving any fossil traces (Hauser et al., 2002; Lieberman, 1984). The
322 comparative approach therefore constitutes a promising method to circumvent this problem
323 (Hauser et al., 2002) and accordingly, much effort has been put into investigating
324 combinatorial capacities in primates (Zuberbühler, 2018). Although studies on our closest
325 living relatives can provide insight into the phylogenetic origins of linguistic traits, they are
326 less useful for informing what factors may have promoted the emergence of our generative
327 capacity (since features could be either homologue/derived or analogue/independently

328 evolved traits). Expanding the comparative approach to include more distantly related species
329 (e.g. birds and non-primate mammals) can help to investigate and identify convergent
330 evolutionary mechanisms, and hence selective conditions, that drive the emergence of
331 combinatorial abilities.

332 Here, we provided an overview of the current evidence of combinatorial structures in
333 non-human animals and show that sound and call combinations can be found across diverse
334 (avian and mammalian) species. Given the structural diversity combinatoriality can take, we
335 propose that a crude division of vocal structures into combinatoriality/phonology versus
336 compositionality/syntax might be an over-simplification trivialising the intriguing complexity
337 of animal vocal structures (c.f. Zuberbühler (2018)). Ultimately, distinctions and subsequent
338 comparisons of combinatorial mechanisms appear to be central to resolving outstanding
339 questions including i) whether we can reveal universal principles that generally drive
340 combinatoriality, and ii) whether different combinatorial mechanisms might be the result of
341 variations in social and ecological conditions or cognitive capacities.

342 To conclude, there exists a considerable diversity in the types of combinatorial
343 structures produced by non-human animals. Such a patchwork of different combinatorial
344 strategies across the animal kingdom implies that different combinatorial mechanisms can
345 emerge independently. From a language evolution perspective, such a finding might suggest
346 language's generative system also represents an assemblage of individually evolved traits
347 (Townsend et al., 2018), rather than a "package" evolved in a sudden evolutionary transition
348 with no similarities in other species (Bolhuis, Tattersall, Chomsky, & Berwick, 2014).
349 Accordingly, with more data on animal combinatoriality (including insights on their
350 distribution, diversity, and the underlying computational processes) a systematic comparative
351 approach may eventually generate intriguing insights into the evolution of communication
352 systems, and ultimately the emergence of language's generative system.

353

354 Acknowledgments

355 We thank Marta Manser for discussions and Zinaida Bogdanova for illustrations.

356

357 References

- 358 ¹ Antunes, R., Schulz, T., Gero, S., Whitehead, H., Gordon, J., & Rendell, L. (2011).
359 Individually distinctive acoustic features in sperm whale codas. *Animal Behaviour*, *81*(4),
360 723-730. doi:10.1016/j.anbehav.2010.12.019
- 361 ² Arnold, K., & Zuberbühler, K. (2006). Language evolution: Semantic combinations in
362 primate calls. *Nature*, *441*(7091), 303. doi:10.1038/441303a
- 363 ³ Arnold, K., & Zuberbühler, K. (2008). Meaningful call combinations in a non-human
364 primate. *Current Biology*, *18*(5), R202-R203. doi:10.1016/j.cub.2008.01.040
- 365 ⁴ Arnold, K., & Zuberbühler, K. (2012). Call combinations in monkeys: Compositional or
366 idiomatic expressions? *Brain and Language*, *120*(3), 303-309.
367 doi:10.1016/j.bandl.2011.10.001
- 368 ⁵ Berwick, R. C., Okanoya, K., Beckers, G. I. J. L., & Bolhuis, J. J. (2011). Songs to syntax:
369 the linguistics of birdsong. *Trends in Cognitive Sciences*, *15*(3), 113-121.
370 doi:10.1016/j.tics.2011.01.002
- 371 ⁶ Blumstein, D. T. (2007). The Evolution, Function, and Meaning of Marmot Alarm
372 Communication. *Advances in the Study of Behavior*, *37*, 371-401. doi:10.1016/s0065-
373 3454(07)37008-3
- 374 ⁷ Bohn, K. M., Schmidt-French, B., Ma, S. T., & Pollak, G. D. (2008). Syllable acoustics,
375 temporal patterns, and call composition vary with behavioral context in Mexican free-
376 tailed bats. *Journal of the Acoustical Society of America*, *124*(3), 1838-1848.
377 doi:10.1121/1.2953314
- 378 ⁸ Bohn, K. M., Smarsh, G. C., & Smotherman, M. (2013). Social context evokes rapid
379 changes in bat song syntax. *Animal Behaviour*, *85*(6), 1485-1491.
380 doi:10.1016/j.anbehav.2013.04.002
- 381 ⁹ Bolhuis, J. J., Tattersall, I., Chomsky, N., & Berwick, R. C. (2014). How Could Language
382 Have Evolved? *PloS Biology*, *12*(8), e1001934. doi:10.1371/journal.pbio.100193
- 383 ¹⁰ Bowling, D. L., & Fitch, W. T. (2015). Do Animal Communication Systems Have
384 Phonemes? *Trends in Cognitive Sciences*, *19*(10), 555-557. doi:10.1016/j.tics.2015.08.011
- 385 ¹¹ Briefer, E. F., Rybak, F., & Aubin, T. (2013). Does true syntax or simple auditory object
386 support the role of skylark song dialect? *Animal Behaviour*, *86*(6), 1131-1137.
387 doi:10.1016/j.anbehav.2013.09.019
- 388 ¹² Candiotti, A., Zuberbühler, K., & Lemasson, A. (2012). Context-related call combinations
389 in female Diana monkeys. *Animal Cognition*, *15*(3), 327-339. doi:10.1007/s10071-011-
390 0456-8
- 391 ¹³ Cäsar, C., Byrne, R., Young, R. J., & Zuberbühler, K. (2012). The alarm call system of
392 wild black-fronted titi monkeys, *Callicebus nigrifrons*. *Behavioral Ecology and*
393 *Sociobiology*, *66*(5), 653-667. doi:10.1007/s00265-011-1313-0

- 394 ¹⁴ Casär, C., Zuberbühler, K., Young, R. J., & Byrne, R. W. (2013). Titi monkey call
395 sequences vary with predator location and type. *Biology Letters*, *9*(5), 20130535.
396 doi:10.1098/rsbl.2013.0535
- 397 ¹⁵ Catchpole, C. K., & Slater, T. L. B. (1995). *Bird Song: Themes and variations*. New York:
398 Cambridge University Press.
- 399 ¹⁶ Cheney, D. L., & Seyfarth, R. M. (1990). *How monkeys see the world*. Chicago: The
400 University of Chicago Press.
- 401 ¹⁷ Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- 402 ¹⁸ Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper &
403 Row.
- 404 ¹⁹ Christiansen, M. H., & Kirby, S. (2003). Language evolution: consensus and
405 controversies. *Trends in Cognitive Sciences*, *7*(7), 300-307. doi:10.1016/s1364-
406 6613(03)00136-0
- 407 ²⁰ Clarke, E., Reichard, U. H., & Zuberbühler, K. (2006). The Syntax and Meaning of Wild
408 Gibbon Songs. *PLoS ONE*, *1*(1), e73. doi:10.1371/journal.pone.0000073
- 409 ²¹ Clay, Z., & Zuberbühler, K. (2009). Food-associated calling sequences in bonobos.
410 *Animal Behaviour*, *77*(6), 1387-1396. doi:10.1016/j.anbehav.2009.02.016
- 411 ²² Clay, Z., & Zuberbühler, K. (2011). Bonobos extract meaning from call sequences. *PLoS*
412 *ONE*, *6*(4), e18786. doi:10.1371/journal.pone.0018786
- 413 ²³ Clucas, B. A., Freeberg, T. M., & Lucas, J. Y. R. (2004). Chick-a-dee call syntax, social
414 context, and season affect vocal responses of Carolina chickadees (*Poecile carolinensis*).
415 *Behavioral Ecology and Sociobiology*, *57*(2), 187-196. doi:10.1007/s00265-004-0847-9
- 416 ²⁴ Collier, K., Bickel, B., van Schaik, C. P., Manser, M. B., & Townsend, S. W. (2014).
417 Language evolution: syntax before phonology? *Proceedings of the Royal Society B:*
418 *Biological Sciences*, *281*(1788), 20140263. doi:10.1098/rspb.2014.0263
- 419 ²⁵ Collier, K., Townsend, S. W., & Manser, M. B. (2017). Call concatenation in wild
420 meerkats. *Animal Behaviour*, *134*, 257-269. doi:10.1016/j.anbehav.2016.12.014
- 421 ²⁶ Courter, J. R., & Ritchison, G. (2010). Alarm calls of tufted titmice convey information
422 about predator size and threat. *Behavioral Ecology*, *21*(5), 936-942.
423 doi:10.1093/Beheco/Arq086
- 424 ²⁷ Coye, C., Ouattara, K., Arlet, M. E., Lemmasson, A., & Zuberbühler, K. (2018). Flexible
425 use of simple and combined calls in female Campbell's monkey. *Animal Behaviour*, *141*,
426 171-181. doi:10.1016/j.anbehav.2018.05.014
- 427 ²⁸ Coye, C., Ouattara, K., Zuberbühler, K., & Lemasson, A. (2015). Suffixation influences
428 receivers' behaviour in non-human primates. *Proceedings of the Royal Society B:*
429 *Biological Sciences*, *282*(1807), 20150265. doi:10.1098/rspb.2015.0265
- 430 ²⁹ Coye, C., Zuberbühler, K., & Lemasson, A. (2016). Morphologically structured
431 vocalizations in female Diana monkeys. *Animal Behaviour*, *115*, 97-105.
432 doi:10.1016/j.anbehav.2016.03.010
- 433 ³⁰ Crockford, C., & Boesch, C. (2005). Call combinations in wild chimpanzees. *Behaviour*,
434 *142*(4), 397-421. doi:10.1163/1568539054012047
- 435 ³¹ Darwin, C. (1871). *The descent of man and selection in relation to sex*. New York: D.
436 Appleton and Company.
- 437 ³² Deaux, E. C., Allen, A. P., Clarke, J. A., & Charrier, I. (2016). Concatenation of 'alert' and
438 'identity' segments in dingoes' alarm calls. *Scientific Reports*, *6*, 30556.
439 doi:10.1038/srep30556

- 440 ³³ DuBois, A. L., Nowicki, S., & Searcy, W. A. (2009). Swamp sparrows modulate vocal
441 performance in an aggressive context. *Biology Letters*, *5*(2), 163-165.
442 doi:10.1098/rsbl.2008.0626
- 443 ³⁴ Engesser, S., Crane, J. M., Savage, J. L., Russell, A. F., & Townsend, S. W. (2015).
444 Experimental Evidence for Phonemic Contrasts in a Nonhuman Vocal System. *PloS*
445 *Biology*, *13*(6), e1002171. doi:10.1371/journal.pbio.1002171
- 446 ³⁵ Engesser, S., Ridley, A. R., Manser, M. B., Manser, A., & Townsend, S. W. (2018).
447 Internal acoustic structuring in pied babbler recruitment cries specifies the form of
448 recruitment. *Behavioral Ecology*. doi:10.1093/beheco/ary088
- 449 ³⁶ Engesser, S., Ridley, A. R., & Townsend, S. W. (2016). Meaningful call combinations and
450 compositional processing in the southern pied babbler. *Proceedings of the National*
451 *Academy of Sciences*, *113*(21), 5976-5981. doi:10.1073/pnas.1600970113
- 452 ³⁷ Engesser, S., Ridley, A. R., & Townsend, S. W. (2017). Element repetition rates encode
453 functionally distinct information in pied babbler 'clucks' and 'purrs'. *Animal Cognition*,
454 *20*(5), 953-960. doi:10.1007/s10071-017-1114-6
- 455 ³⁸ Fedurek, P., Zuberbühler, K., & Dahl, C. D. (2016). Sequential information in a great ape
456 utterance. *Scientific Reports*, *6*, 38226. doi:10.1038/srep38226
- 457 ³⁹ Ficken, M. S., Hailman, E. D., & Hailman, J. P. (1994). The Chick-a-Dee Call System of
458 the Mexican Chickadee. *The Condor*, *96*(1), 70-82. doi:10.2307/1369065
- 459 ⁴⁰ Fitch, W. T. (2018). What animals can teach us about human language: the phonological
460 continuity hypothesis. *Current Opinion in Behavioral Sciences*, *21*, 68-75.
461 doi:10.1016/j.cobeha.2018.01.014
- 462 ⁴¹ Freeberg, T. M., & Lucas, J. R. (2002). Receivers respond differently to chick-a-dee calls
463 varying in note composition in Carolina chickadees, *Poecile carolinensis*. *Animal*
464 *Behaviour*, *63*(5), 837-845. doi:10.1006/anbe.2001.1981
- 465 ⁴² Freeberg, T. M., & Lucas, J. R. (2012). Information Theoretical Approaches to Chick-a-
466 dee Calls of Carolina Chickadees (*Poecile carolinensis*). *Journal of Comparative*
467 *Psychology*, *126*(1), 68-81. doi:10.1037/A0024906
- 468 ⁴³ Geissmann, T. (2002). Duet-splitting and the evolution of gibbon songs. *Biological*
469 *Reviews*, *77*(1), 57-76. doi:10.1017/S1464793101005826
- 470 ⁴⁴ Hailman, J. P. (1989). The Organization of Major Vocalizations in the Paridae. *The Wilson*
471 *Bulletin*, *101*(2), 305-343.
- 472 ⁴⁵ Hailman, J. P., & Ficken, M. S. (1986). Combinatorial animal communication with
473 computable syntax: Chick-a-dee calling qualifies as 'Language' by structural linguistics.
474 *Animal Behaviour*, *34*(6), 1899-1901. doi:10.1016/s0003-3472(86)80279-2
- 475 ⁴⁶ Hailman, J. P., Ficken, M. S., & Ficken, R. W. (1985). The Chick-a-Dee Calls of *Parus*
476 *atricapillus* - a Recombinant System of Animal Communication Compared with Written-
477 English. *Semiotica*, *56*(3-4), 191-224. doi:10.1515/Semi.1985.56.3-4.191
- 478 ⁴⁷ Hauser, M. D. (1997). *The Evolution of Communication*. Cambridge: MIT Press.
- 479 ⁴⁸ Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The Faculty of Language: What Is It,
480 Who Has It, and How Did It Evolve? *Science*, *298*(5598), 1569-1579.
481 doi:10.1126/science.298.5598.1569
- 482 ⁴⁹ Hedwig, D., Mundry, R., Robbins, M. M., & Boesch, C. (2014). Contextual correlates of
483 syntactic variation in mountain and western gorilla close-distance vocalizations:
484 Indications for lexical or phonological syntax? *Animal Cognition*, *18*(2), 423-435.
485 doi:10.1007/s10071-014-0812-6

- 486 ⁵⁰ Hockett, C. F. (1960). The Origin of Speech. *Scientific American*, 203, 88-111.
- 487 ⁵¹ Holland, J., Dabelsteen, T., & Paris, A. L. (2000). Coding in the song of the wren:
488 importance of rhythmicity, syntax and element structure. *Animal Behaviour*, 60(4), 463-
489 470. doi:10.1006/anbe.2000.1529
- 490 ⁵² Humboldt, W. v. (1999). *On Language*. Translated by Peter Heath. Cambridge:
491 Cambridge University Press.
- 492 ⁵³ Hurford, J. (2007). *The origins of meaning*. Oxford: Oxford University Press.
- 493 ⁵⁴ Hurford, J. (2012a). Linguistics from an Evolutionary Point of View. In R. Kempson, M.
494 Fernando, & N. Ashe (Eds.), *Handbook of the Philosophy of Science: Linguistics* (Vol. 14,
495 pp. 473-498). Amsterdam: Elsevier Science.
- 496 ⁵⁵ Hurford, J. (2012b). *The origins of grammar*. Oxford: Oxford University Press.
- 497 ⁵⁶ Janik, V., & Slater, P. (1997). Vocal Learning in Mammals. *Advances in the Study of*
498 *Behavior*, 26, 59-99. doi:10.1016/s0065-3454(08)60377-0
- 499 ⁵⁷ Jansen, D. A., Cant, M. A., & Manser, M. B. (2012). Segmental concatenation of
500 individual signatures and context cues in banded mongoose (*Mungos mungo*) close calls.
501 *BMC Biol*, 10, 97. doi:10.1186/1741-7007-10-97
- 502 ⁵⁸ Kershenbaum, A., Ilany, A., Blaustein, L., & Geffen, E. (2012). Syntactic structure and
503 geographical dialects in the songs of male rock hyraxes. *Proceedings of the Royal Society*
504 *B: Biological Sciences*, 279(1740), 1-8. doi:10.1098/rspb.2012.0322
- 505 ⁵⁹ King, S. L., & Janik, V. M. (2013). Bottlenose dolphins can use learned vocal labels to
506 address each other. *Proceedings of the National Academy of Sciences*, 110(32), 13216-
507 13221. doi:10.1073/pnas.1304459110
- 508 ⁶⁰ Lemasson, A., Ouattara, K., Bouchet, H., & Zuberbühler, K. (2010). Speed of call delivery
509 is related to context and caller identity in Campbell's monkey males. *Naturwissenschaften*,
510 97(11), 1023-1027. doi:10.1007/s00114-010-0715-6
- 511 ⁶¹ Lieberman, P. (1984). *The Biology and Evolution of Language*. Cambridge, MA: Harvard
512 University Press.
- 513 ⁶² Mahurin, E. J., & Freeberg, T. M. (2008). Chick-a-dee call variation in Carolina
514 chickadees and recruiting flockmates to food. *Behavioral Ecology*, 20(1), 111-116.
515 doi:10.1093/beheco/arn121
- 516 ⁶³ Manser, M. B. (2001). The acoustic structure of suricates' alarm calls varies with predator
517 type and the level of response urgency. *Proceedings of the Royal Society B: Biological*
518 *Sciences*, 268(1483), 2315-2324. doi:10.1098/rspb.2001.1773
- 519 ⁶⁴ Marler, P. (1977). *The structure of animal communication sounds. Recognition of complex*
520 *acoustic signals: report of Dahlem workshop*. Berlin: Abakon Verlagsgesellschaft.
- 521 ⁶⁵ Marler, P. (2000). Origins of Music and Speech: Insights from Animals. In W. L. Nils, B.
522 Merker, & S. Brown (Eds.), *The Origins of Music* (pp. 31-48). Cambridge, MA: The MIT
523 Press.
- 524 ⁶⁶ Marler, P., Dufty, A., & Pickert, R. (1986). Vocal communication in the domestic chicken:
525 II. Is a sender sensitive to the presence and nature of a receiver? *Animal Behaviour*, 34(1),
526 194-198. doi:10.1016/0003-3472(86)90023-0
- 527 ⁶⁷ Marler, P., & Slabbekoorn, H. (1999). *Nature's Music: The science of birdsong*. New
528 York: Academic Press.
- 529 ⁶⁸ Mitani, J. C., & Marler, P. (1989). A Phonological Analysis of Male Gibbon Singing
530 Behavior. *Behaviour*, 109(1/2), 20-45.

- 531 ⁶⁹ Nousek, A. E., Slater, P. J., Wang, C., & Miller, P. J. (2006). The influence of social
532 affiliation on individual vocal signatures of northern resident killer whales (*Orcinus orca*).
533 *Biology Letters*, 2(4), 481-484. doi:10.1098/rsbl.2006.0517
- 534 ⁷⁰ Nowak, M. A. (2000). Evolutionary biology of language. *Philosophical Transactions of*
535 *The Royal Society B: Biological Sciences*, 355(1403), 1615-1622.
536 doi:10.1098/rstb.2000.0723
- 537 ⁷¹ Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009a). Campbell's monkeys concatenate
538 vocalizations into context-specific call sequences. *Proceedings of the National Academy of*
539 *Sciences*, 106(51), 22026-22031. doi:10.1073/pnas.0908118106
- 540 ⁷² Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009b). Campbell's Monkeys Use
541 Affixation to Alter Call Meaning. *PLoS ONE*, 4(11), e7808.
542 doi:10.1371/journal.pone.0007808
- 543 ⁷³ Payne, R. S., & McVay, S. (1971). Songs of Humpback Whales. *Science*, 173(3997), 585-
544 597. doi:10.1126/science.173.3997.585
- 545 ⁷⁴ Pinker, S., & Jackendoff, R. (2005). The faculty of language: what's special about it?
546 *Cognition*, 95(2), 201-236. doi:10.1016/j.cognition.2004.08.004
- 547 ⁷⁵ Radford, A. N., & Ridley, A. R. (2006). Recruitment Calling: A Novel Form of Extended
548 Parental Care in an Altricial Species. *Current Biology*, 16(17), 1700-1704.
549 doi:10.1016/j.cub.2006.06.053
- 550 ⁷⁶ Rek, P., & Osiejuk, T. S. (2012). Temporal patterns of broadcast calls in the corncrake
551 encode information arbitrarily. *Behavioral Ecology*, 24(2), 547-552.
552 doi:10.1093/beheco/ars196
- 553 ⁷⁷ Rendall, D. (2013). Q&A: Cognitive ethology - inside the minds of other species. *BMC*
554 *Biol*, 11(108), 1-6. doi:10.1186/1741-7007-11-108
- 555 ⁷⁸ Robinson, J. G. (1984). Syntactic Structures in the Vocalizations of Wedge-Capped
556 Capuchin Monkeys, *Cebus olivaceus*. *Behaviour*, 90(1), 47-79.
- 557 ⁷⁹ Rohrmeier, M., Zuidema, W., Wiggins, G. A., & Scharff, C. (2015). Principles of structure
558 building in music, language and animal song. *Philosophical Transactions of The Royal*
559 *Society B: Biological Sciences*, 370(1664), 20140097. doi:10.1098/rstb.2014.0097
- 560 ⁸⁰ Schel, A. M., Candiotti, A., & Zuberbühler, K. (2010). Predator-detering alarm call
561 sequences in Guereza colobus monkeys are meaningful to conspecifics. *Animal*
562 *Behaviour*, 80(5), 799-808. doi:10.1016/j.anbehav.2010.07.012
- 563 ⁸¹ Schel, A. M., Tranquilli, S., & Zuberbühler, K. (2009). The alarm call system of two
564 species of black-and-white colobus monkeys (*Colobus polykomos* and *Colobus guereza*).
565 *Journal of Comparative Psychology*, 123(2), 136-150. doi:10.1037/a0014280
- 566 ⁸² Schlenker, P., Chemla, E., Arnold, K., Lemasson, A., Ouattara, K., Keenan, S., Stephan,
567 C., Ryder, R., & Zuberbühler, K. (2014). Monkey semantics: two 'dialects' of Campbell's
568 monkey alarm calls. *Linguistics and Philosophy*, 37(6), 439-501. doi:10.1007/s10988-014-
569 9155-7
- 570 ⁸³ Schlenker, P., Chemla, E., Cäsar, C., Ryder, R., & Zuberbühler, K. (2016). Titi semantics:
571 Context and meaning in Titi monkey call sequences. *Natural Language & Linguistic*
572 *Theory*, 35(1), 271-298. doi:10.1007/s11049-016-9337-9
- 573 ⁸⁴ Schlenker, P., Chemla, E., & Zuberbühler, K. (2016). What Do Monkey Calls Mean?
574 *Trends in Cognitive Sciences*, 20(12), 894-904. doi:10.1016/j.tics.2016.10.004
- 575 ⁸⁵ Schulz, T. M., Whitehead, H., Gero, S., & Rendell, L. (2011). Individual vocal production
576 in a sperm whale (*Physeter macrocephalus*) social unit. *Marine Mammal Science*, 27(1),
577 149-166. doi:10.1111/j.1748-7692.2010.00399.x

- 578 ⁸⁶ Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervet monkey alarm calls: Semantic
579 communication in a free-ranging primate. *Animal Behaviour*, 28(4), 1070-1094.
580 doi:10.1016/s0003-3472(80)80097-2
- 581 ⁸⁷ Sievers, C., & Gruber, T. (2016). Reference in human and non-human primate
582 communication: What does it take to refer? *Animal Cognition*, 19(4), 759-768.
583 doi:10.1007/s10071-016-0974-5
- 584 ⁸⁸ Sieving, K. E., Hetrick, S. A., & Avery, M. L. (2010). The versatility of graded acoustic
585 measures in classification of predation threats by the tufted titmouse *Baeolophus bicolor*:
586 exploring a mixed framework for threat communication. *Oikos*, 119(2), 264-276.
587 doi:10.1111/j.1600-0706.2009.17682.x
- 588 ⁸⁹ Suzuki, T. N. (2013). Communication about predator type by a bird using discrete, graded
589 and combinatorial variation in alarm calls. *Animal Behaviour*, 87, 59-65.
590 doi:10.1016/j.anbehav.2013.10.009
- 591 ⁹⁰ Suzuki, T. N. (2016). Semantic communication in birds: evidence from field research over
592 the past two decades. *Ecol Res*, 31(3), 307-319. doi:10.1007/s11284-016-1339-x
- 593 ⁹¹ Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2016). Experimental evidence for
594 compositional syntax in bird calls. *Nature Communication*, 7, 10986.
595 doi:10.1038/ncomms10986
- 596 ⁹² Templeton, C. N., Greene, E., & Davis, K. (2005). Allometry of Alarm Calls: Black-
597 Capped Chickadees Encode Information About Predator Size. *Science*, 308(5730), 1934-
598 1937. doi:10.1126/science.1108841
- 599 ⁹³ ten Cate, C. (2017). Assessing the uniqueness of language: Animal grammatical abilities
600 take center stage. *Psychonomic Bulletin and Review*, 24(1), 91-96. doi:10.3758/s13423-
601 016-1091-9
- 602 ⁹⁴ Townsend, S. W., Engesser, S., Stoll, S., Zuberbühler, K., & Bickel, B. (2018).
603 Compositionality in animals and humans. *PloS Biology*, 18(8), e2006425.
604 doi:10.1371/journal.pbio.2006425
- 605 ⁹⁵ Townsend, S. W., Koski, S. E., Byrne, R. W., Slocombe, K. E., Bickel, B., Boeckle, M.,
606 Goncalves, I. B., Burkart, J. M., Flower, T., Gaunet, F., et al. (2016). Exorcising Grice's
607 ghost: an empirical approach to studying intentional communication in animals. *Biological*
608 *Reviews*, 92(3), 1427-1433. doi:10.1111/brv.12289
- 609 ⁹⁶ Townsend, S. W., & Manser, M. B. (2013). Functionally Referential Communication in
610 Mammals: The Past, Present and the Future. *Ethology*, 119(1), 1-11.
611 doi:10.1111/eth.12015
- 612 ⁹⁷ Wheatcroft, D. (2015). Repetition rate of calls used in multiple contexts communicates
613 presence of predators to nestlings and adult birds. *Animal Behaviour*, 103, 35-44.
614 doi:10.1016/j.anbehav.2015.02.009
- 615 ⁹⁸ Wheeler, B. C., & Fischer, J. (2012). Functionally referential signals: a promising
616 paradigm whose time has passed. *Evolutionary Anthropology*, 21(5), 195-205.
617 doi:10.1002/evan.21319
- 618 ⁹⁹ Yip, M. J. (2006). The search for phonology in other species. *Trends in Cognitive*
619 *Sciences*, 10(10), 442-446. doi:10.1016/J.Tics.2006.08.001
- 620 ¹⁰⁰ Zuberbühler, K. (2008). Audience effects. *Current Biology*, 18(5), R189-R190.
621 doi:10.1016/j.cub.2007.12.041
- 622 ¹⁰¹ Zuberbühler, K. (2018). Combinatorial capacities in primates. *Current Opinion in*
623 *Behavioral Sciences*, 21, 164-169. doi:10.1016/j.cobeha.2018.03.015

624 ¹⁰² Zuidema, W., & de Boer, B. (2018). The evolution of combinatorial structure in language.
625 *Current Opinion in Behavioral Sciences*, *21*, 138-144. doi:10.1016/j.cobeha.2018.04.011
626