



**HAL**  
open science

## Sound categorization by crocodilians

Julie Thévenet, Mounia Kehy, Nicolas Boyer, Aurélie Pradeau, Léo Papet,  
Etienne Gaudrain, Nicolas Grimault, Nicolas Mathevon

► **To cite this version:**

Julie Thévenet, Mounia Kehy, Nicolas Boyer, Aurélie Pradeau, Léo Papet, et al.. Sound categorization by crocodilians. *iScience*, inPress, 10.1016/j.isci.2023.106441 . hal-04041406

**HAL Id: hal-04041406**

**<https://hal.science/hal-04041406>**

Submitted on 22 Mar 2023

**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.

# Journal Pre-proof



Sound categorization by crocodilians

Julie Thévenet, Mounia Kehy, Nicolas Boyer, Aurélie Pradeau, Léo Papet, Etienne Gaudrain, Nicolas Grimault, Nicolas Mathevon

PII: S2589-0042(23)00518-7

DOI: <https://doi.org/10.1016/j.isci.2023.106441>

Reference: ISCI 106441

To appear in: *ISCIENCE*

Received Date: 25 November 2022

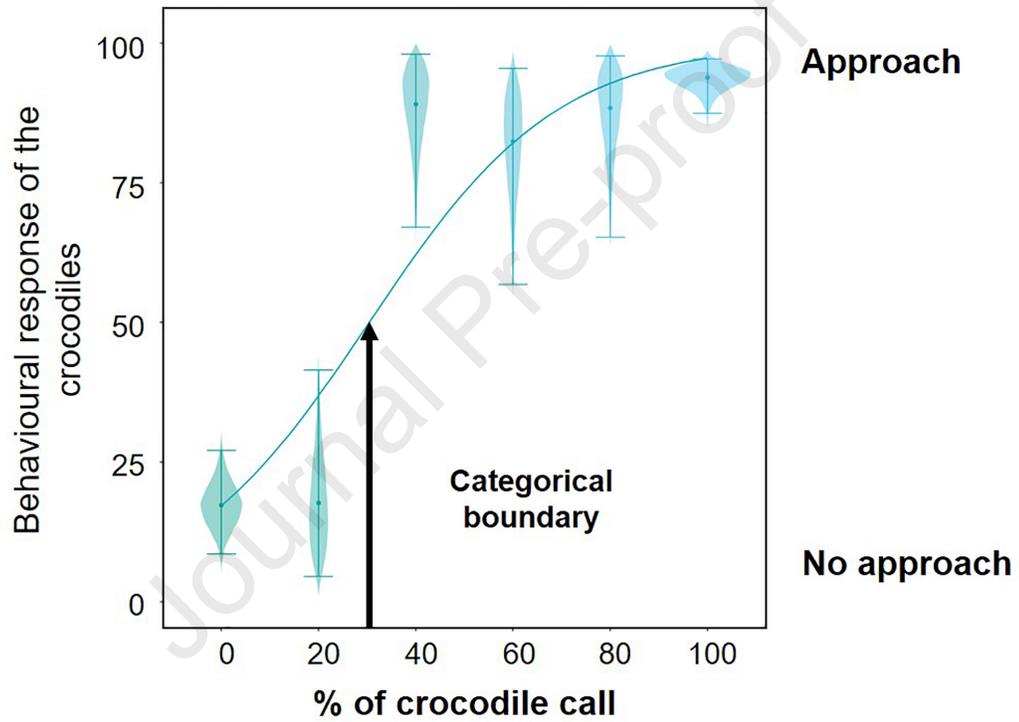
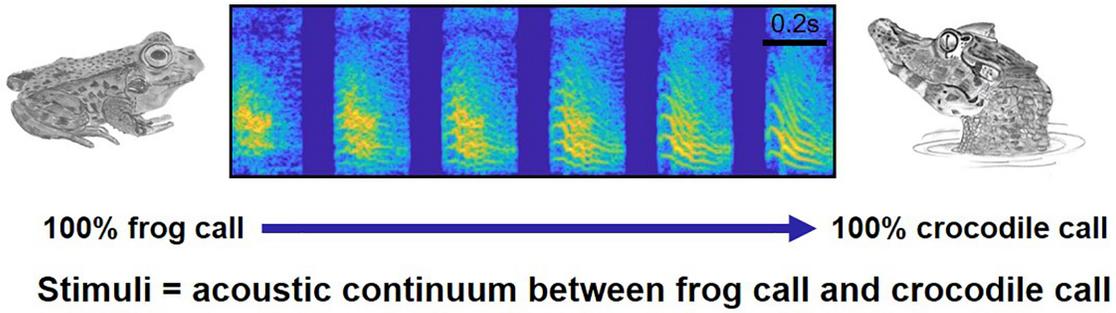
Revised Date: 7 February 2023

Accepted Date: 14 March 2023

Please cite this article as: Thévenet, J., Kehy, M., Boyer, N., Pradeau, A., Papet, L., Gaudrain, E., Grimault, N., Mathevon, N., Sound categorization by crocodilians, *ISCIENCE* (2023), doi: <https://doi.org/10.1016/j.isci.2023.106441>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2023



# 1 Sound categorization by crocodilians

2 **Julie Thévenet<sup>1,2\*</sup>, Mounia Kehy<sup>1</sup>, Nicolas Boyer<sup>1</sup>, Aurélie Pradeau<sup>1</sup>, Léo Papet<sup>1</sup>, Etienne**  
3 **Gaudrain<sup>2</sup>, Nicolas Grimault<sup>2‡</sup>, Nicolas Mathevon<sup>1,3‡</sup>**

4  
5 <sup>1</sup>ENES Bioacoustics Research Laboratory, CRNL, CNRS, Inserm, University of Saint-Etienne, Saint-  
6 Etienne 42100, France;

7 <sup>2</sup>Equipe Cognition Auditive et Psychoacoustique, CRNL, CNRS, Inserm, University Lyon 1, Bron 69500,  
8 France;

9 <sup>3</sup>Institut universitaire de France

10

11 ‡ These authors contributed equally

12 \*Lead contact and author for correspondence (julie.thevenet.phd@gmail.com)

13

## 14 **SUMMARY**

15 Rapidly sorting the information contained in a stream of stimuli is a major challenge for  
16 animals. One cognitive mechanism for achieving this goal is categorization, where the  
17 receiving individual considers a continuous variation of a stimulus as belonging to discrete  
18 categories. Using playback experiments in a naturalistic setup, here we show that crocodiles  
19 confronted with an acoustic continuum ranging from a frog call to a crocodile call classify each  
20 acoustic variant into one of these two categories, establishing a meaningful boundary where  
21 no acoustic boundary exists. With GO/NOGO experiments, we then observe that this  
22 boundary is defined along the continuum following learning. We further demonstrate that  
23 crocodilians rely on the spectral envelope of sounds to categorize stimuli. This study suggests  
24 that sound categorization in crocodilians is a pre-wired faculty allowing rapid decision making,  
25 and highlights the learning-dependent plasticity involved in defining the boundary between  
26 sound categories.

27

## 28 **KEYWORDS**

29 Auditory perception, categorization, learning, GO/NOGO operant conditioning, animal  
30 communication, crocodiles.

## 31 INTRODUCTION

32 Detecting and quickly processing information from the massive flow of information emanating  
33 from the auditory scene is a major challenge for animals <sup>1</sup>. Although an auditory stimulus can  
34 be accurately described physically, its perception by animals depends primarily on their  
35 auditory abilities and perceptual processes <sup>2-5</sup>. Being able to categorize stimuli can help  
36 process auditory information and make decisions about how to respond. In this process,  
37 animals integrate continuous variation in a sound stimulus as belonging to discrete categories  
38 carrying different information. This results in a peak in discrimination between stimuli of  
39 different categories versus stimuli of the same category <sup>6-8</sup>. However, categorizing sounds,  
40 i.e., deciding which stimulus corresponds to which object in the sound scene, such as a  
41 conspecific call rather than a call from another species, can be complicated especially when  
42 the two signals are acoustically close. As suggested by the signal detection theory <sup>9</sup>, the  
43 optimal decision in terms of the probability of responding to a conspecific call rather than a  
44 heterospecific call depends on the relative probability of each of the two species being present  
45 at the location of the sound source, the acoustic cues carried by the sound stimulus, and the  
46 cost and benefit of the decision made by the animal. An acoustic feature threshold value - or  
47 informational threshold- is identified, determining the point at which animals decide on their  
48 categorical classification and behavioral response <sup>10,11</sup>.

49 Previous works have shown that frequency parameters are often used to quickly classify  
50 sounds into categories. The cricket *Teleogryllus oceanii* categorizes pulsed sounds into  
51 attractive versus repulsive when they cross a frequency threshold of 16 kHz <sup>12</sup>. The great tits  
52 *Parus major* bases its classification of song notes on their fundamental frequency <sup>13</sup>. The  
53 japanese macaque *Macaca fuscata* relies on the temporal position of a frequency peak to  
54 classify the contact calls of its species into two functional categories <sup>14</sup>. Categorical thresholds  
55 may also be related to the temporal characteristics of sound stimuli. Thus, the duration of  
56 song notes and calls is the criterion used to establish the threshold between two informational  
57 categories and guide the behavioral response of the swamp sparrow *Melospiza georgiana* <sup>15</sup>  
58 and the mouse *Mus domesticus* <sup>16</sup> respectively. However, our understanding of sound  
59 categorization processes in animals remains limited since, among tetrapods, only a few  
60 species have been studied <sup>17-19</sup>. In the present paper, we test whether crocodylians are able  
61 to form functional categories from continuous sets of stimuli.

62 Crocodylians belong to Archosaurs, a monophyletic group that includes the modern birds as  
63 well as past dinosaurs and pterosaurs. As such, they share behaviors with other archosaurs  
64 such as forms of parental care, since the crocodylian parents (female, male or both depending  
65 on the crocodylian species) often protect the nest and young from predators. While they are  
66 more distantly related to mammals, crocodylians exhibit many ecological traits that have  
67 evolved convergently with this group, especially with mammalian apex predators <sup>20</sup>. The  
68 unique position of crocodylians in the tree of life thus puts these animals in a privileged  
69 position for comparative research. Understanding how crocodylians hear the world is key to  
70 decipher the evolutionary history of auditory perception in all vertebrates.

71 Crocodylians use their hearing abilities both to identify potential prey and during their social  
72 interactions. The auditory communication channel is essential in their early years where calls  
73 are used by the young to communicate with each other and with their parent <sup>21</sup>. Although the  
74 sensory organs of crocodylians are quite well known, our knowledge about how they perceive  
75 and process the sounds they hear remains scarce. Their position as top predators and their  
76 auditory abilities suggest that crocodylians are super-efficient and quick to analyze their sound  
77 scene, derive relevant information, and make appropriate decisions in terms of behavioral  
78 response <sup>22</sup>. For example, crocodiles are able to accurately localize a sound source <sup>23-25</sup> and  
79 use this ability to effectively perform spatial unmasking to identify a relevant sound despite  
80 background noise <sup>26</sup>. In this work, we test the hypothesis that crocodylians are able to partition  
81 an acoustic continuum into categories and identify the acoustic features used for this  
82 categorization. Being able to form meaningful categories from auditory scenes should allow  
83 crocodylians to facilitate rapid decision making, and adjust a behavioral response such as an  
84 approach to the sound source accordingly.

85 In a first experiment, we test whether this ability to partition an acoustic continuum into  
86 discontinuous categories is innate by observing the response of naive crocodiles to frog and  
87 crocodile calls as well as to chimeric signals containing a variable proportion of frog and  
88 crocodile acoustic features. Second, we test whether the boundary between these categories  
89 can be modulated by learning by training caimans to approach a speaker emitting a crocodile  
90 call and ignore the frog call (GO/NOGO procedure) and then testing them with a continuum  
91 of acoustic chimeras. Finally, we investigate the acoustic basis of this categorization using  
92 acoustic chimeras based on only certain orthogonal dimensions of the signal ( $f_0$ , aperiodicity,  
93 spectral envelope). Our results establish that crocodylians follow behavioral decision rules

94 based on the establishment of sound categories constructed from the spectral envelope of  
95 signals.

96

## 97 **RESULTS**

### 98 **Experiment 1: Naïve crocodilians respond discontinuously to continuous stimuli.**

99 We tested juvenile Nile crocodiles *Crocodylus niloticus* ( $n = 14$ ) with an acoustic continuum  
100 between two signals that, while sharing some acoustic similarities, came from two different  
101 animal species and thus had different biological meanings: a Saharan frog call *Pelophylax*  
102 *saharicus*, a species routinely participating in the spring night sound scene of the zoo where  
103 we did the experiments, and a contact call of a young Nile crocodile, which previous studies  
104 have shown to be attractive to crocodiles<sup>27</sup>. The frog call and the crocodile call are harmonic  
105 series in the same frequency range (1000-3500 Hz and 250-4000 Hz respectively) and are of  
106 comparable duration (175 ms and 200 ms respectively). To create the acoustic continuum  
107 (acoustic morphing), we first decomposed the frog and crocodile calls into three orthogonal  
108 acoustic dimensions: the pitch ( $f_0$ ), the aperiodicity (ap) and the spectral envelope (env). Each  
109 acoustic chimera was then re-synthesized by combining the three dimensions of a frog call  
110 and those of a crocodile call in different proportions (see Methods for details). Finally, the  
111 acoustic continuum consisted of 11 calls, ranging from a 100% frog call to a 100% crocodile  
112 call, through 9 chimeric calls where the relative proportion of the three dimensions between  
113 the two species varied from 90% frog / 10% crocodile to 10% frog / 90% crocodile, with a 10%  
114 step between each signal (Figure 1A).

115 The experiment took place in an experimental basin (6 X 7 meters), with four loudspeakers  
116 placed near the water on the edge of the basin at four opposite locations (Figure 1B). From  
117 sunset until the end of the experiment, the 4 speakers emitted frog calls, in an unsynchronized  
118 manner (i.e. every 9, 10, 11 and 13 seconds for speakers n°1 to n°4 respectively), to mimic a  
119 frog chorus to which crocodiles are used to. As the experiment took place during fall, no  
120 natural frog calls interfered with the artificial chorus. The tested crocodile was then exposed  
121 to a succession of experimental tests using each acoustic chimera of the continuum (see Figure  
122 1B for a schematic of the experimental setup).

123 As illustrated in Figure 1C, the crocodiles were attracted to the speaker when the stimulus had  
124 a very high proportion of crocodile features. Specifically, their response score increased

125 sharply for the 10% frog / 90% crocodile stimulus whereas it was low for stimuli with a lower  
126 proportion of crocodile features (response score 1.2 point higher for the 90% crocodile signal  
127 compared to the 80% crocodile signal, 95% CI [0.1, 2.3], 98.5% of the posterior distribution is  
128 positive).

129

130 **Experiments 2 and 3: Crocodylians trained to distinguish two sounds establish a sharp**  
131 **boundary along their acoustic continuum.**

132 The ability of crocodylians to categorize sound signals forming an acoustic continuum between  
133 a "frog call" and a "crocodile call" was further tested in the laboratory via a GO-NOGO  
134 experiment (Experiment 2) where young dwarf Cuvier's caiman *Paleosuchus palpebrosus* (n =  
135 5) were trained to respond positively (i.e. by approaching the speaker) to a crocodile call. The  
136 stimuli used were identical to those used in Experiment 1 except that we reduced the step  
137 between each signal: the continuum consisted of 6 calls, from a 100% frog call to a 100%  
138 crocodile call, through 4 chimeric calls where the relative proportion of the three dimensions  
139 between the two species varied from 80% frog / 20% crocodile to 20% frog / 80% crocodile,  
140 with a 20% step between each signal.

141 As shown in Figure 2A, the caimans learned to systematically move toward the speaker  
142 emitting the 100% crocodile signal (GO signal). At the beginning of training, the response rate  
143 was around 67.4%, while it was only 38.3% for the 100% frog signal. After 2 weeks of Go/No-  
144 Go training, this response rate settled above 75% and the caimans were 76.2% more likely to  
145 respond to the GO stimulus than to the NOGO stimulus (95% CI [66.5, 84.1]).

146 The results of Experiment 2 are shown in Figure 2B. Caimans did not respond to the 100% frog  
147 / 0% crocodile signal, and their response to the 80% frog / 20% crocodile signal was not  
148 significantly higher (0.6% higher, 95% CI [-14.7, 23.4]). In contrast, their response increased  
149 sharply from the 60% frog / 40% crocodile chimeric signal (69.6% more likely to respond with  
150 the 40% crocodile signal compared to the 20% crocodile signal, 95% CI [41.4, 87.8]). This high  
151 level of response was maintained for all the following signals on the acoustic continuum, with  
152 no marked difference from the response to the 60% frog / 40% crocodile signal (probability in  
153 obtaining a stronger response: with the 40% frog / 60% crocodile signal = -6.3%, 95% CI [-32.1,  
154 17.0]; with the 20% frog / 80% crocodile signal = -0.6, 95% CI [-24.1, 21.8]); with the 0% frog /  
155 100% crocodile signal = 4.7%, 95% CI [-5.6, 25.4]).

156 The response of the caimans thus appears clearly dichotomous, with a boundary between two  
157 signal categories located between the 80% frog / 20% crocodile signal and the 60% frog / 40%  
158 crocodile signal (probability of response 69.2% higher for the second category, 95% CI [55.0,  
159 79.5]). This boundary was remarkably stable across individuals: all animals tested reached a  
160 maximum response rate of 33% for the 80% frog / 20% crocodile signal, whereas they all  
161 reached a response rate of at least 75% for the 60% frog / 40% crocodile signal. These results  
162 were confirmed when comparing the latency to respond to the stimulus between the frog  
163 category (100% and 80% frog stimuli, mean latency of  $12.8 \pm 12.3$  s) and the crocodile category  
164 (40% crocodile and above stimuli, mean latency of  $5.3 \pm 6.4$  s): the latency decreased by 8.9 s  
165 (95% CI [-25.5, -1.3]), while there was no difference of latency for stimuli within the frog  
166 category (difference of -2.0 s, 95% CI [-14.2, 23.9] between the 100% and 80% frog stimuli)  
167 and for stimuli within the crocodile category (difference of -1.6 s between the 40% and 60%  
168 crocodile stimuli, 95% CI [-16.1, 12.0]; -2.9 s between the 40% and 80% crocodile stimuli, 95%  
169 CI [-17.9, 6.8]; and 0.5 s between the 40% and 100% crocodile stimuli, 95% CI [-14.7, 5.6]).

170 After completing Experiment 2, the same caimans were tested in another experimental  
171 paradigm, also based on the GO-NOGO approach, but in a two-choice testing procedure  
172 (alternative choice test). The goal was to test the ability of crocodilians to categorize acoustic  
173 signals when confronted with two sound sources alternately emitting a pair of neighboring  
174 acoustic chimeras within the continuum. 5 pairs of stimuli were formed: 100% frog / 0%  
175 crocodile with 80% frog / 20% crocodile, 80% frog / 20% crocodile with 60% frog / 40%  
176 crocodile, 60% frog / 40% crocodile with 40% frog / 60% crocodile, 40% frog / 60% crocodile  
177 with 20% frog / 80% crocodile, and 20% frog / 80% crocodile with 0% frog / 100% crocodile.

178 The results obtained in Experiment 3 (Figure 2C) strengthen the conclusions drawn from  
179 Experiment 2. Indeed, when presented with a choice between the 80% frog / 20% crocodile  
180 signal and the 60% frog / 40% crocodile signal, the caimans showed a clear preference for the  
181 second signal: they were 40.5% more likely to choose the 40% crocodile signal (95% CI [18.5,  
182 58.7]). Conversely, when confronted with the 100% frog / 0% crocodile and 80% frog / 20%  
183 crocodile signal pair, the caimans responded only weakly and equally to both signals  
184 (probability of 65.0 to get no response, 95% CI [53.3, 75.2], difference of only 2.3%, 95% CI [-  
185 12.3, 16.4] to choose one or another stimulus). For all pairs involving signals containing at least  
186 40% crocodile acoustic features, the caimans showed no preference, responding strongly to  
187 either signal in the pair (pair [40%, 60%]: 87.2% of response, 95% CI [78.5, 93.3], probability

188 of a different response between the two signals = 13.0%, 95% CI [-12.4, 37.2]; pair [60%, 80%]:  
 189 83.0% (95% CI [73.7, 90.2]), probability of a different response = 4.3%, 95% CI [-20.2, 28.1];  
 190 pair [80%, 100%]: 80.9%, 95% CI [71.2, 88.6], probability of a different response = -1.4%, 95%  
 191 CI [-25.3, 22.2]. The latencies to respond were however not different from one pair of stimuli  
 192 to another (minor difference of 2.8 s, 95% CI [-1.4, 8.8] between the between category pair  
 193 [20%, 40%] and the frog category pair [0%, 20%]; -1.5 s, 95% CI [-4.4, 0.6] of difference  
 194 between the pair [20%, 40%] and the crocodile category pairs).

195

#### 196 **Experiment 4: Sound categorization is based on the spectral envelope.**

197 To investigate which signal dimensions ( $f_0$ , ap, and/or env) are used by crocodilians to  
 198 categorize calls as "frog" or "crocodile", we first tested the caimans with acoustic chimeras  
 199 bearing on only one or two dimensions. In this first step, each chimera was constructed based  
 200 on a 100% frog / 0% crocodile call, and one or two dimensions were raised to its value in the  
 201 100% crocodile call. In addition to the two control calls 100% frog / 0% crocodile and 0% frog  
 202 / 100% crocodile, the following 6 acoustic chimeras were tested:

- 203 - 0%  $f_0$ , 100% ap, 100% env frog / 100%  $f_0$ , 0% ap, 0% env crocodile;
- 204 - 100%  $f_0$ , 0% ap, 100% env frog / 0%  $f_0$ , 100% ap, 0% env crocodile;
- 205 - 100%  $f_0$ , 100% ap, 0% env frog / 0%  $f_0$ , 0% ap, 100% env crocodile;
- 206 - 0%  $f_0$ , 0% ap, 100% env frog / 100%  $f_0$ , 100% ap, 0% env crocodile;
- 207 - 0%  $f_0$ , 100% ap, 0% env frog / 100%  $f_0$ , 0% ap, 100% env crocodile;
- 208 - 100%  $f_0$ , 0% ap, 0% env frog / 100%  $f_0$ , 100% ap, 100% env crocodile.

209 Only the acoustic chimeras containing the spectral envelope of the crocodile call induced a  
 210 significant behavioral response from the caimans (Table 1, Figure 3A). The presence of the  
 211 crocodile spectral envelope alone was sufficient to induce a response identical to that  
 212 obtained with the 0% frog / 100% crocodile signal (negligible difference of response  
 213 probability of -8.3%, 95% CI [32, 3.9]). We tend to obtain concordant results if we compare  
 214 the latency times (Table 2).

215 Comparison of labelling between the control acoustic continuum and the acoustic continuum  
 216 involving only the spectral envelope shows that caimans establish the boundary between the  
 217 two categories "frog" and "crocodile" at the same location on both continua (67.7% frog /

218 32.3% crocodile and 67% frog / 33% crocodile, respectively; Figure 3B). Furthermore, there is  
219 no significant difference in behavioral response between full chimeras and chimeras involving  
220 only the spectral envelope throughout the acoustic continua (median difference of 5.1 %, 95%  
221 CI [-4.6, 15.8], Table 3). Consistent results are obtained if latency times are compared (Table  
222 3).

223

## 224 **DISCUSSION**

225 In this study, we demonstrate through playback experiments that young crocodiles  
226 confronted with a continuum of acoustic chimeras ranging from a frog call to a crocodile call  
227 classify each acoustic variant into one or the other of these two categories, thus establishing  
228 a clear meaningful boundary where no acoustic boundary yet exists. By conducting GO/NOGO  
229 experiments, we suggest that this boundary between two categories can be established along  
230 the acoustic continuum by learning. Finally, we show that crocodilians' categorization of  
231 acoustic stimuli is based primarily on the spectral envelope rather than the pitch or aperiodic  
232 dimension of the sound signals.

233 Confronted with chimeric signals consisting of a mixture of "frog" and "crocodile"  
234 characteristics, the tested individuals made the choice to approach or not approach the  
235 speaker. One would have expected variations in behavioural response to be linearly correlated  
236 with the percentage of "crocodile" signal in the chimeras. This was not the case. While our  
237 animals had previously heard frog calls, which are naturally present in the zoo ponds, as well  
238 as calls from conspecifics, they had never experienced acoustic chimeras. Our results thus  
239 suggest that naive crocodilians practice sound categorization innately.

240 It is likely that crocodilians can form categories under other circumstances. Because  
241 crocodilians have developed a vocal repertoire, it is possible that this ability to categorize is  
242 also useful during acoustic communication between individuals. For example, young  
243 crocodilians emit cries with "graded" acoustic characteristics by playing on intensity, energy  
244 distribution in the frequency spectrum and frequency modulation. A study conducted with  
245 the black caiman *Melanosuchus niger* showed that the repertoire of young caimans extends  
246 along an acoustic continuum from contact calls allowing the cohesion of the group of young  
247 to distress calls inducing maternal protection<sup>28</sup>. It is possible -although not yet demonstrated-  
248 that young caiman and their mothers establish two distinct categories along the contact-

249 distress acoustic continuum, which would facilitate decision making. However, it remains to  
250 be established where the boundary between contact and distress lies. This boundary is likely  
251 to vary according to the context. If the level of risk experienced by the young (in terms of  
252 predation pressure for example) is high, it is likely that the boundary between the contact and  
253 distress categories may be at a different level than if the level of risk is lower. This hypothesis  
254 remains to be tested.

255 The labelling experiment (Experiment 2) may suggest a role of learning in defining the  
256 boundary between two sound categories. It may seem surprising that crocodilians trained to  
257 respond to a 0% frog-100% crocodile signal have a low categorization threshold (when the  
258 signal is "33% crocodile"). This result contrasts with that obtained in the first experiment,  
259 where the crocodiles received no prior training and were never rewarded. Although it should  
260 be kept in mind that Experiments 1 and 2 are hardly comparable (different crocodilian species,  
261 different experimental conditions), one might still have expected trained individuals to have  
262 a high response threshold, since only the "100% crocodile" signal led to a reward. In our  
263 laboratory conditioning protocol, animals were not confronted with any aversive reinforcer,  
264 meaning that there is no risky choice<sup>9</sup>. This may have led individuals to never take the risk of  
265 missing a potentially rewarded signal, and thus be more likely to respond to signals with only  
266 a small proportion of crocodile features. In animals, context-dependent decisions are indeed  
267 strongly related to estimates of the probabilities of making a good choice<sup>29,30</sup>.

268 The alternative choice experiment (Experiment 3) confirms the results obtained in the  
269 labelling experiment (Experiment 2). Tested individuals expressed a particularly distinct choice  
270 when offered the 80% frog / 20% crocodile and 60% frog / 40% crocodile signals. However,  
271 for all other pairs of signals, they chose randomly between the two acoustic chimera.

272 What are the mechanisms underlying this sound categorization process? In particular, can it  
273 be considered as categorical perception? Categorical perception is defined as the process by  
274 which stimuli are considered similar or different depending on whether or not they fall into  
275 the same perceptual category<sup>7,19</sup>. Strictly speaking, two features characterize a categorical  
276 perception: 1) the labelling of stimuli in one or the other of the categories considered, which  
277 corresponds to what we observed in our experiments with crocodilians; 2) a "category  
278 boundary effect" in which two stimuli belonging to the same category are differentiated less  
279 rapidly than two stimuli located on either side of the category boundary. We could not  
280 rigorously test this second condition with our crocodiles. Testing the category boundary effect

281 would require numerous GO/NOGO tests (aiming to discriminate the two stimuli in each pair  
282 of stimuli) that are not reasonably feasible with these animals due to the small number of  
283 individuals available in the lab combined with a long learning time (unlike what is possible with  
284 humans, birds or monkeys)<sup>13,14,31-37</sup>. In our protocol, it was also difficult to use a  
285 habituation/dishabituation approach<sup>15,12,38</sup> because crocodylians become satiated with food  
286 rewards rather quickly and quickly lose motivation to respond. Despite these experimental  
287 limitations, our observations strongly suggest the presence of a categorical perception of  
288 sound stimuli in the crocodylians tested. Indeed, while caimans were conditioned to respond  
289 to the 0% frog - 100% crocodile stimulus in the choice experiment (Experiment 3), they only  
290 showed a preference for the signal containing the most crocodile trait only when the pair "80%  
291 frog / 20% crocodile and 60% frog / 40% crocodile" was emitted. For the other combinations  
292 (40/60, 60/80, and 80/100 crocodile traits), test subjects showed no preference for either of  
293 the two signals in each pair (Figure 2C). In other words, the crocodylians tested reacted as if  
294 they were truly unable to perceptually differentiate between these two stimuli.

295 Our final experiment questions the acoustic basis for the categorization performed by  
296 crocodylians. It turns out that crocodylians rely on the spectral envelope of sound signals to  
297 categorize stimuli while they seem to ignore the pitch and aperiodic component of the signals.  
298 Using the spectral envelope to identify a complex sound made up of a series of harmonics is a  
299 widespread and reliable strategy<sup>37</sup>. In human perception, the spectral envelope of the voice  
300 (i.e the voice timbre) is related to the vocal tract length of the speaker which is related to the  
301 speaker height, weight, age and gender<sup>39</sup>. As such, this feature is often used to encode  
302 speaker identity<sup>40</sup>. The spectral envelope, corresponding to the timbre of the emitter, is  
303 indeed an invariant acoustic parameter, while the  $f_0$  contour and the aperiodicity might vary  
304 with the context<sup>41</sup>. The spectral envelope would be a more stable and reliable cue to  
305 categorize crocodile calls, all types of calls combined.

306 In conclusion, our study shows that crocodylians exhibit an ability to categorize sound stimuli.  
307 This ability certainly contributes to an awareness of auditory scenes in their everyday life. It  
308 complements other processes such as their ability to locate precisely the position of a sound  
309 source<sup>23,42,24,25</sup> and their ability to identify a meaningful sound despite background noise by  
310 spatial unmasking<sup>26</sup>. Given their lifestyle, crocodylians are confronted with complex sound  
311 environments, and must be able to make quick decisions. Being able to quickly categorize  
312 sounds according to their functional relevance is certainly extremely helpful.

### 313 **LIMITATIONS OF STUDY**

314 One limitation of this study is that we had to work with two different species of crocodilians.  
315 This makes it difficult to compare results between experiment 1 (performed with Nile  
316 crocodiles) and the other experiments (performed with Cuvier's caimans).

317 Another point is that we did not have the possibility to rigorously test the "category boundary  
318 effect". While our study clearly demonstrates that crocodilians respond in a categorized  
319 manner to an acoustic continuum, we cannot prove that the mechanism of this categorization  
320 is strictly a "categorical perception" as defined in the literature. As mentioned above, testing  
321 "categorical perception" would have required the implementation of numerous GO/NOGO  
322 tests, which is impractical given the complexity of working with these animals.

323

### 324 **ACKNOWLEDGEMENTS**

325 The authors are grateful to "Crocoparc" zoo in Agadir, Morocco (Luc Fougérol, Ariane  
326 Marinetti, Leila Sdigui, Philippe and Christine Alleon, and the staff), to Olivier Marquis for  
327 providing us with the caimans and to David Reby for his advice. This research has been funded  
328 by the Institut universitaire de France (NM), the Labex CeLyA (Lyon Center of Acoustics ANR-  
329 10-LABX-60, PhD funding to Julie Thévenet), the CNRS and the University of Saint-Etienne.

330

### 331 **AUTHORS CONTRIBUTIONS**

332 JT: conceptualization, data acquisition, analysis, writing; MK: data acquisition, analysis; EG:  
333 synthesis of chimeric sound signals; NB: data acquisition, manipulation and well-being of the  
334 animals; AP: well-being of the animals; LP: conceptualization, data acquisition; NG:  
335 conceptualization, data acquisition, writing; NM: conceptualization, data acquisition, writing.

336

### 337 **DECLARATION OF INTERESTS**

338 The authors declare no competing interests.

339

340

341 **MAIN FIGURES**342 **Figure 1. Categorization tests with naive crocodilians reveal that they switch their behavioral**  
343 **response across a specific acoustic boundary along a signal continuum (Experiment 1).**

344 (A) Acoustic continuum of frog/crocodile chimeric signals, ranging from a 'frog' signal (100% frog / 0%  
345 croc) to a 'crocodile' signal (0% frog / 100% croc). The 'frog' signal was used to create a frog chorus.  
346 Only 5 of the 10 signals tested are shown here (the entire continuum spanned from a 100% frog / 0%  
347 crocodile signal to a 0% frog / 100% crocodile signal, with 10% increments; see Methods for details of  
348 the acoustic morphing procedure).

349 (B) Experimental set-up. Throughout each experiment (average duration = 2.5 hours), all four speakers  
350 were emitting 100% frog calls (1 call every 9 to 13 seconds, different rhythm for each speaker,  
351 reproducing a frog chorus). During each test, one of the speakers switched to one of the chimeric  
352 signals (duration of a test = 4 minutes; tests separated by at least 10 minutes; different test speaker  
353 from one test to another; see Methods for details). The crocodile's response to the chimeric signal was  
354 assessed for 10 minutes after the start of each test. Each individual (N = 14 Nile crocodiles) was tested  
355 successively with all chimeric signals, presented in random order during the experiment.

356 (C) Behavioral response of crocodiles to chimeric signals. Solid circles show the fitted probabilities of  
357 getting a response (the crocodile approached the test speaker) expressed as the median of posterior  
358 distribution with 95% CI. Violin plots show the distribution of the fitted values for each signal. The  
359 abrupt increase in approaches to the test speaker from the 10% frog / 90% croc signal (marked by the  
360 black arrow) suggests sound categorization. This result was obtained with naive, untrained crocodiles,  
361 which had therefore never heard the test signals before the experiment.

362

363 **Figure 2. Crocodilians trained to respond to a meaningful signal highlight their ability to form a**  
364 **sensory boundary along an acoustic continuum (Experiments 2 and 3).**

365 (A) Training procedure: Young caimans (N = 5) were trained using food reward to approach the speaker  
366 only when it emits a 100% croc signal (GO-NOGO procedure). As the graph on the right shows, training  
367 reinforces the natural propensity of animals to respond to that signal.

368 (B) Experiment 2: Categorization tests along the acoustic continuum suggest a sensory boundary  
369 between the 20% and 40% croc signals (black arrow). Solid circles show the fitted probabilities of  
370 getting a response (the crocodile approached the test speaker) expressed as the median of posterior  
371 distribution with 95% CI. Violin plots show the distribution of the fitted values for each signal.

372 (C) Experiment 3: Fitted probability of choosing between two chimeric signals in an alternative-choice  
373 task (upper graph), and difference between the probabilities measured for each pair of signals (below).  
374 Choice between neighboring chimeric signals is effective only between the 20 and 40% croc signals  
375 (black arrow).

376

377 **Figure 3. The categorization of frog/croc chimeric signals by crocodilians is based on the spectral**  
378 **envelope of the sound stimulus, excluding other acoustic dimensions (Experiment 4).**

379 (A) Results of playback experiments with animals trained to respond to crocodile call. Solid circles show  
380 the fitted probabilities of getting a response (the crocodile approached the test speaker) expressed as  
381 the median of posterior distribution with 95% CI. Violin plots show the distribution of the fitted values  
382 for each signal. One-dimensional and two-dimensional chimeric stimuli were created from a 100% frog

383 call by transforming into 100% croc three acoustic dimensions separately or in pairs: pitch ( $f_0$ ),  
384 aperiodic dimension (ap), and spectral envelope (env), and their combinations  $f_0 + ap$ ,  $f_0 + env$ , and  
385  $env + ap$ . Only stimuli with 100% crocodile spectral envelope elicited a behavioral response. The  
386 subjects did not respond to stimuli with only 100%  $f_0$  croc or 100% ap croc.

387 (B) When the spectral envelope is the only dimension to change in the acoustic continuum (in pink),  
388 the sensory boundary remains close to that identified for the original continuum where all three  
389 dimensions of the signal change together (boundaries between the 32 and 33% croc signals: blue and  
390 pink arrows).

391

392

Journal Pre-proof

393 **MAIN TABLES**

394

Stimuli	Fitted response probability (%)	Median [95% CI] difference with the NOGO stimulus (%)	Median [95% CI] difference with the GO stimulus (%)
<b>100% frog / 0% crocodile</b>	15.8		-77.8 [-85.1, -68.7]
<b>0% frog / 100% crocodile</b>	93.9	77.8 [68.7, 85.1]	
<b>100% <math>f_0</math> crocodile</b>	5.5	-9.7 [-20.5, 8.1]	-87.8 [-94.8, -69.8]
<b>100% env crocodile</b>	85.2	68.8 [45.6, 82.8]	-8.6 [-32, 3.9]
<b>100% ap crocodile</b>	5.5	-9.6 [-20.5, 8.0]	-87.8 [-94.7, -70.0]
<b>100% <math>f_0</math> + env crocodile</b>	85.4	69.0 [45.4, 83.1]	-8.3 [-31.5, 4.0]
<b>100% <math>f_0</math> + ap crocodile</b>	5.5	-9.6 [-20.4, 8.5]	-87.8 [-94.7, -69.5]
<b>100% env + ap crocodile</b>	78.7	62.4 [37.3, 79.1]	-15.0 [-40.3, 0.7]

395

396 **Table 1. Effect of the stimulus type on the caiman behavioral reaction (Experiment 4).** Medians of  
 397 posterior distribution (%) and 95% CI.

398

Stimuli	Fitted latency (sec)	Median [95% CI] difference with the NOGO stimulus (sec)	Median [95% CI] difference with the GO stimulus (sec)
100% frog / 0% crocodile	17.0		10.3 [4.5, 21.4]
0% frog / 100% crocodile	6.7	-10.3 [-21.4, -4.5]	
100% $f_0$ crocodile	20.7	3.0 [-11.1, 35.8]	13.3 [1.5, 49.4]
100% env crocodile	6.5	-10.2 [-21.9, -2.6]	-0.1 [-4.3, 7.0]
100% ap crocodile	14.0	-2.8 [-14.6, 14.0]	7.2 [-0.4, 27.1]
100% $f_0$ + env crocodile	6.8	-9.9 [-21.8, -2.1]	0.1 [-4.1, 7.4]
100% $f_0$ + ap crocodile	7.9	-8.7 [-20.5, 0.7]	1.2 [-3.5, 11.5]
100% env + ap crocodile	12.5	-4.4 [-15.5, 8.7]	5.8 [-0.5, 21]

399

400 **Table 2. Effect of the stimulus type on the caiman latency time to react (Experiment 4).** Medians of  
 401 posterior distribution and 95% CI.

402

Stimuli	Fitted response probability (%)		Fitted latency (sec)	
	Full morphing	Partial morphing	Full morphing	Partial morphing
<b>15% crocodile</b>	9.5 [2.0, 28.3]	6.3 [1.3, 21]	12.3 [6.0, 26.5]	17.5 [8.4, 37.1]
<b>20% crocodile</b>	3.9 [0.2, 24.4]	2.6 [0.1, 17.5]	11.9 [5.9, 24.7]	16.9 [8.5, 34.5]
<b>25% crocodile</b>	7.1 [0.2, 46.6]	4.7 [0.1, 37.0]	9.9 [4.9, 19.9]	14.1 [6.8, 29.4]
<b>30% crocodile</b>	19.8 [2.4, 57.5]	13.6 [1.6, 47.2]	10.9 [5.3, 22.5]	15.5 [7.6, 31.5]
<b>35% crocodile</b>	60.6 [25.1, 86.4]	49.9 [17.9, 80.3]	10.6 [5.3, 21.5]	15.0 [7.6, 30.7]
<b>40% crocodile</b>	80.6 [41.7, 95.6]	73.0 [31.6, 93.1]	8.0 [4.0, 16.1]	11.4 [5.8, 22.8]

403

404 **Table 3. Effect of stimulus type on the caiman behavioral reaction and latency time (Experiment 4).**

405 Medians of posterior distribution (%) and 95% CI.

406

407

## 408 **STAR METHODS**

### 409 **RESOURCE AVAILABILITY**

#### 410 **Lead contact**

411 Further information and requests for resources should be directed to and will be fulfilled by  
412 the lead contact, Julie Thévenet (julie.thevenet.phd@gmail.com).

#### 413 **Materials availability**

414 This study did not generate new unique reagents.

#### 415 **Data and code availability**

- 416 • Dataset have been deposited at Zenodo and are publicly available as of the date of  
417 publication<sup>43</sup>. The DOI is listed in the key resources table.
- 418 • All original statistical code has been deposited at Zenodo and is publicly available as  
419 of the date of publication<sup>43</sup>. The DOI is listed in the key resources table.
- 420 • Any additional information required to reanalyze the data reported in this paper is  
421 available from the lead contact upon request.

422

## 423 **EXPERIMENTAL MODEL AND SUBJECT DETAILS**

### 424 **Experiment 1: Labelling experiment with untrained naïve crocodilians**

425 This experiment was performed at *Crocoparc* zoo (Agadir, Morocco). The animals (n = 14 Nile  
426 crocodiles *Crocodylus niloticus*, 1-2 month old, sex unknown) were housed together in an  
427 outdoor enclosure, not visible to the public. They had never been included in any experiment  
428 before. All studies were conducted under the ethical approval of the ENES laboratory (n° D  
429 42-218-0901) and complied with Moroccan and French laws.

430

### 431 **Experiment 2: Labelling experiments with trained crocodilians**

432 The experiment was conducted at the ENES laboratory. We worked with young dwarf Cuvier's  
433 caiman *Paleosuchus palpebrosus*, born in captivity at the "Zoo de Paris" (N = 5, including 2  
434 females and 3 males; age: 2.5 years; size: 59 ± 2 cm). These animals had never been included  
435 in an experimental protocol prior to the experiment. Individuals were housed all together in a

436 large enclosure (16 m<sup>2</sup>). The room was maintained at  $28 \pm 1$  °C, with a light on from 9:00 am  
437 to 9:00 pm.

438

### 439 **Experiment 3: Alternative-choice test with trained crocodilians**

440 We performed this experiment at the ENES laboratory, with the same dwarf Cuvier's  
441 caimans *Paleosuchus palpebrosus* as in experiment 2 (n = 5 individuals).

442

### 443 **Experiment 4: Identification of acoustic cues supporting category labelling**

444 This experiment was conducted at the ENES laboratory, and involved the same dwarf  
445 Cuvier's caimans (n = 5 individuals), after they had performed experiments 2 and 3.

446

## 447 **METHOD DETAILS**

### 448 **Experiment 1: Labelling experiment with untrained naïve crocodilians**

449 To limit pseudo-replication, we used 4 frog calls (recorded by Léo Papet in the zoo) and 4  
450 juvenile crocodile calls (recorded from free-ranging young crocodiles during a field expedition  
451 in the Okavango Delta by N. Mathevon and T. Aubin). Each of the 4 acoustic *continua* used  
452 during the experiment was made from one of these frog calls and one of these crocodile calls.

453 Signal decomposition in three orthogonal dimensions ( $f_0$ , aperiodicity, spectral envelope) was  
454 performed with the open source Matlab toolbox STRAIGHT<sup>44,45</sup>. The  $f_0$  contour represents the  
455 pitch variations as a function of time. The spectral envelope shape and scale are related to the  
456 shape and size of the vocal tract, or in general, of the resonating cavity used for  
457 vocal production. The aperiodicity matrix indicates where in time and frequency the extracted  
458  $f_0$  applies, i.e. it is used to represent breathiness in the vocalization. Due to some limitation in  
459 the parameters' variation, STRAIGHT failed to extract the  $f_0$  contour of the frog calls. To do so,  
460 we used an auto-correlation to estimate the modulation frequency of the temporal envelope  
461 in successive temporal windows. This homemade code is provided as an open source resource.  
462 The acoustic chimera and original calls were then re-synthesized using STRAIGHT by  
463 combining the three orthogonal dimensions of a frog call and those of a crocodile call in  
464 different proportions. The aperiodicity and the spectral envelope dimensions from frog and  
465 crocodile calls were linearly combined as provided by STRAIGHT, while their  $f_0$  value were first

466 log-transformed before being linearly combined, and then transformed back to a linear Hertz  
467 scale.

468 To avoid habituation, each individual was tested in only one experimental session (one night),  
469 with a single set of acoustic chimeras. Three hours before the start of the experiment, the  
470 crocodile was released into the experimental basin. The experiments, conducted in the fall,  
471 started one hour after sunset as it is during the night that the frogs sing the most during spring  
472 and the crocodiles are the most active. All sound signals used in the experiment were emitted  
473 at an intensity of 67 dBA at 1 m from the speaker (FoxPro Fusion© with Visaton SL 87 ND  
474 internal speaker; see Figure S1B and Figure S1C for technical specifications).

475 At the time of a test, a loudspeaker was randomly selected among the four, eliminating the  
476 one or ones from which the crocodile was closest. For 2 minutes, this test speaker continued  
477 to emit frog calls, but with a rhythm of 1 every 12 seconds. The choice of a fixed rate of  
478 emission, the same for all tests of all crocodiles, eliminated the risk of obtaining different  
479 behavioral responses following different rates of emission. At the end of the first 2 minutes,  
480 the test speaker started to emit for the next 2 minutes one of the acoustic chimeras (chosen  
481 at random between the 90% frog / 10% crocodile and the 0% frog / 100% crocodile) at the  
482 same rate of one call every 12 seconds. At the end of the 2 minutes, the test speaker resumed  
483 emitting frog calls at its original rate. The next test was initiated at least 10 minutes later, from  
484 a different speaker (stimulus chosen at random). The behavior of the crocodile was observed  
485 and evaluated blindly and independently by 4 observers for 10 minutes from the start of the  
486 test. The entire experiment was filmed by an infrared camera (ABUS TVCC34010).

487 The intensity of the crocodiles' behavioral response to sound stimuli was scored from 0 to 5  
488 according to the following scale: 0 = no response; 1 = set in motion without swimming; 2 =  
489 swimming without preferred direction; 3 = swimming in the direction of the speaker but not  
490 exceeding 1/3 of the distance to the speaker; 4 = swimming in the direction of the speaker  
491 without approaching the speaker within 50 centimeters; 5 = swimming in the direction of the  
492 speaker and approaching within 50 centimeters.

493

## 494 **Experiment 2: Labelling experiments with trained crocodilians**

495 The experiment was conducted in a square-shaped pool (1.75 m wide), placed in a soundproof  
496 cabin, with a water depth of 12 cm allowing the animals to swim freely (water temperature:

497 29°C). Four speakers (Audiopro Bravo Allroom Sat, see Figure S1E and Figure S1F for technical  
498 specifications) were placed just above the water surface at the four corners of the pool (Figure  
499 2A). Sound playback was controlled by the experimenter from outside the booth using a  
500 computer delivering the signals to an amplifier (Yamaha AX-397) connected to the speakers.  
501 A remote-controlled reward system was associated with each speaker, which allowed for  
502 multiple delivery of food rewards during a single test session. All tests were performed in low  
503 light conditions to limit the stress of the animals and increase their motivation to respond  
504 (these animals are more active at night). The whole experiment was filmed (ABUS TVCC34010  
505 camera), which allowed real time and delayed observations.

506 Training took place in weekly sessions via operant conditioning using a food reward. The  
507 subject was first placed in the pool and did not receive any stimulation for an acclimatization  
508 period of at least 3 min (we always waited until the animal had stopped exploring the pool).  
509 Then, the GO stimulus (sequence of crocodile calls) and the NOGO stimulus (sequence of frog  
510 calls) were presented three to seven times each in a random order (10 calls per sequence at a  
511 rate of one call every  $2.5 \pm 0.4$  sec, total sequence duration = 35 sec;  $10 \pm 3$  sequences in total  
512 per training session). The animal was systematically rewarded if it came within 35 cm of the  
513 active speaker when a sequence of crocodile calls (GO stimulus) was presented. This 35 cm  
514 distance corresponded to the location where the reward (food) was delivered. There was a  
515 minimum delay of 1 min 30 seconds between two trials. The acoustic stimuli were emitted at  
516 a sound level of 60 dB SPL at 1m from the speaker.

517 Once trained, i.e., when the subject responded correctly in more than 75% of the tests, the  
518 subject moved on to weekly test sessions, during which the animal was confronted with three  
519 to five sequences of calls, randomly selected from the acoustic continuum. No rewards were  
520 given during the tests when the crocodile response was measured, regardless of the signal  
521 and the animal's response. However, in order to maintain motivation over the test sessions,  
522 the animal's response was regularly reinforced by repeating the go/no-go protocol, i.e., by  
523 associating a positive response to the GO signal with a food reward. These reinforcements  
524 were randomly distributed during the sessions. During each test session, the caiman heard  $12$   
525  $\pm 3$  sequences of calls, including 40% test stimuli (without reward), 40% GO stimuli, 20% NOGO  
526 stimuli (total duration of a test session = between 20 and 40 minutes). Testing sessions were  
527 repeated until each caiman had been tested three times with each of the test stimuli of the

528 continuum (i.e. total of  $12 \pm 4$  test stimuli per caiman throughout the experiments, with at  
529 least 3 repetitions for each of the 4 test stimuli).

530 The response of the animals to the acoustic stimuli was measured blindly from the videos  
531 using Kinovea software, v0.8.24 Beta <sup>49</sup>, during the 45 seconds following the emission of the  
532 first call of the tested sequence. For each trial, the response was scored as "0" (no approach)  
533 or "1" (reach the food reward system within 35 centimeters of the speaker). We also  
534 measured the latency between the first call in the sequence and the onset of the behavioral  
535 response (first movement), if any.

536

### 537 **Experiment 3: Alternative-choice test with trained crocodilians**

538 The two alternative stimuli were presented antiphonally during 35 seconds through two  
539 speakers located at equal distance from the animal. The overall rhythm of emission was the  
540 same as in the second experiment, which means that only 5 calls were emitted per speaker  
541 instead of 10 calls. A test session always began with a training pair combining a 100% frog /  
542 0% crocodile signal (NOGO) with a 0% frog / 100% crocodile signal (GO). Discrimination tests  
543 only began if the animal moved to the speaker emitting the GO signal twice in a row. If the  
544 animal made the wrong choice, it was trained again with the GO stimulus. Then, the animals  
545 were tested with all the five pairs of stimuli beginning once by one stimuli of the pair and then  
546 the other, in a random order (i.e. 10 pairs of test stimuli were presented by session, total  
547 duration of a test session = between 20 and 40 minutes). In total, each caiman was tested with  
548  $87 \pm 31$  pairs of stimuli throughout the experiment.

549

### 550 **Experiment 4: Identification of acoustic cues supporting category labelling**

551 As in Experiment 2, the stimuli were delivered in a sequence of 10 successive calls (one signal  
552 every  $2.5 \pm 0.4$ ; stimuli presented in a random order). The assessment of caiman behavioral  
553 responses and the statistical approach were identical to those used in Experiment 2. In total,  
554 the caimans were tested with  $18 \pm 3$  test stimuli throughout the experiments.

555 Since the results of this first step showed that the spectral envelope was a decisive acoustic  
556 criterion for categorization, we tested whether the boundary between the two categories  
557 "frog" and "crocodile" was identical between a control continuum where the 3 acoustic  
558 dimensions varied and a continuum where only the spectral envelope varied. Based on the  
559 results of experiments 2 and 3 which indicated that the acoustic boundary should be for a

560 signal containing around 30% crocodile features, we created a series of 6 acoustic chimeras  
561 where the 3 dimensions of the signal varied together, from a 85% frog / 15% crocodile signal  
562 to a 60% frog / 40% crocodile signal, with a 5% step. We also created another series of 6  
563 acoustic chimeras, where only the spectral envelope was involved, from a 100%  $f_0$ , 100% ap,  
564 85% env frog / 0%  $f_0$ , 0% ap, 15% env crocodile signal to a 100%  $f_0$ , 100% ap, 60% env frog /  
565 0%  $f_0$ , 0% ap, 40% env crocodile signal.

566

## 567 **QUANTIFICATION AND STATISTICAL ANALYSIS**

### 568 **Experiment 1: Labelling experiment with untrained naïve crocodilians**

569 Statistical analyses were performed in R (v3.6.2) using Bayesian mixed models fitted with the  
570 brms R package <sup>46</sup>. The advantages of using the Bayesian approach are multiple, including its  
571 high flexibility, quantification of uncertainty in estimates, and intuitive interpretation of  
572 confidence intervals <sup>47,48</sup>. Response scores were modeled using a gaussian function. The model  
573 included subject-specific random intercepts. 5000 iterations were run over four MCMC chains  
574 with the first 500 iterations of each chain used to adjust the algorithm. Because the behavior  
575 of the tested animals was difficult to predict, flat priors were kept for this model. Results were  
576 summarized as medians of the posterior distributions and 95% credible intervals (CIs).  
577 Credible intervals for estimates that do not include zero indicate a credible effect given the  
578 observed data and model structure <sup>48</sup>. Similarly, when contrasting two conditions, CIs  
579 excluding the null value can be inferred to indicate a credible difference between the  
580 conditions.

581

### 582 **Experiment 2: Labelling experiments with trained crocodilians**

583 Response scores were modeled using a logistic function (Bernoulli family) and reaction times  
584 using a shifted-lognormal distribution. All models included subject-specific random intercepts.  
585 5000 iterations were run over four MCMC chains, with the 500 first iterations of each chains  
586 used to tune the algorithm. To improve convergence when modeling response scores, we  
587 specified mildly informative priors for all factor levels (nature of stimulus, normal distribution  
588 of (0, 3)), except for the GO stimulus (set with an informative normal distribution of (2, 2)) and  
589 the NOGO stimulus (normal distribution (-2, 2)). Results were summarized as medians of  
590 posterior distributions and 95% credible intervals (CIs).

591

**592 Experiment 3: Alternative-choice test with trained crocodilians**

593 For each trial, the caiman was considered to have made its choice between the two speakers  
594 when it approached one of them within 45 seconds of the call. The response was scored  
595 blindly. Choices were modeled using a multinomial function (categorical family), with random  
596 intercepts per animal.

597

**598 Experiment 4: Identification of acoustic cues supporting category labelling**

599 The experimental and behavioral response analysis procedures were identical to those in  
600 Experiment 2. In total, the caimans were tested with  $54 \pm 2$  test stimuli throughout the  
601 experiments.

602

603

604

605

606

607

608

609 **KEY RESOURCES TABLE**

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Dataset, acoustic signals and statistical codes	This article	<a href="https://doi.org/10.5281/zenodo.7699715">https://doi.org/10.5281/zenodo.7699715</a>
Experimental models: Organisms/strains		
<i>Crocodylus niloticus</i>	Crocoparc Agadir (Agadir, Morocco)	Taxonomy ID: 8493
<i>Paleosuchus palpebrosus</i>	Parc zoologique de Paris (Paris, France)	Taxonomy ID: 8497
Software and algorithms		
R	R Project	RRID:SCR_001905 <a href="https://www.r-project.org/">https://www.r-project.org/</a>
Matlab 2020b	MathWorks	RRID: SCR_001622 <a href="https://fr.mathworks.com/products/matlab.html">https://fr.mathworks.com/products/matlab.html</a>
Kinovea	Kinovea	<a href="https://www.kinovea.org/">https://www.kinovea.org/</a>

610

611

## 612 REFERENCES

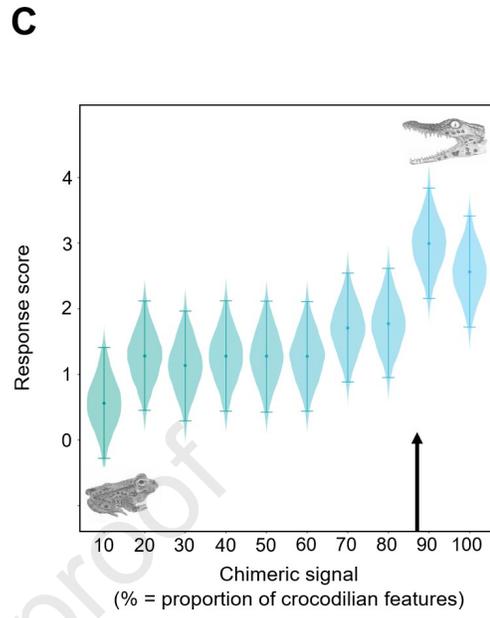
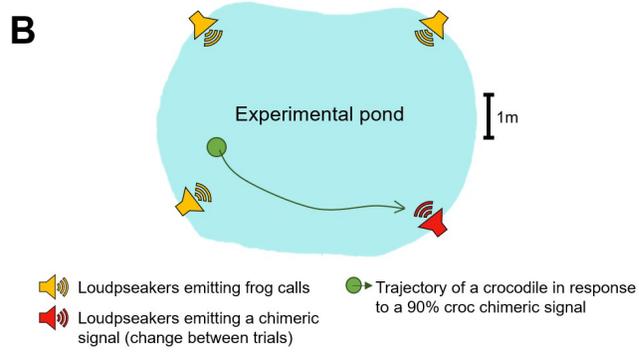
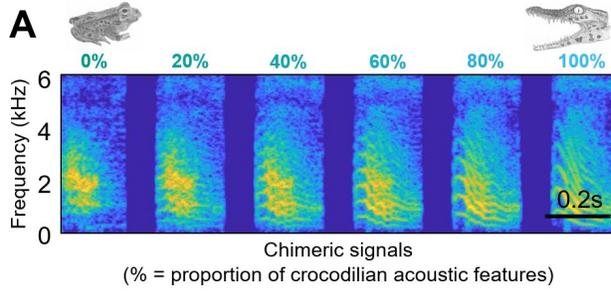
- 613 1. Hulse, S.H. (2002). Auditory scene analysis in animal communication. In *Advances in the Study of*  
614 *Behavior* (Academic Press), pp. 163–200. 10.1016/S0065-3454(02)80008-0.
- 615 2. Guilford, T., and Dawkins, M.S. (1991). Receiver psychology and the evolution of animal signals.  
616 *Animal Behaviour* 42, 1–14. 10.1016/S0003-3472(05)80600-1.
- 617 3. Miller, C.T., and Bee, M.A. (2012). Receiver psychology turns 20: is it time for a broader  
618 approach? *Anim Behav* 83, 331–343. 10.1016/j.anbehav.2011.11.025.
- 619 4. Rowe, C. (2013). Receiver psychology: a receiver’s perspective. *Animal Behaviour* 85, 517–523.  
620 10.1016/j.anbehav.2013.01.004.
- 621 5. Bee, M.A., and Miller, C.T. (2016). Signaler and Receiver Psychology. In *Psychological*  
622 *Mechanisms in Animal Communication Animal Signals and Communication.*, M. A. Bee and C. T.  
623 Miller, eds. (Springer International Publishing), pp. 1–16. 10.1007/978-3-319-48690-1\_1.
- 624 6. Wood, C.C. (1976). Discriminability, response bias, and phoneme categories in discrimination of  
625 voice onset time. *The Journal of the Acoustical Society of America* 60, 1381–1389.  
626 10.1121/1.381231.
- 627 7. Harnad, S. (1987). *Categorical perception: The groundwork of cognition* (Cambridge University  
628 Press).
- 629 8. Caves, E.M., Green, P.A., Zipple, M.N., Peters, S., Johnsen, S., and Nowicki, S. (2018). Categorical  
630 perception of colour signals in a songbird. *Nature* 560, 365–367. 10.1038/s41586-018-0377-7.
- 631 9. Bradbury, J.W., and Vehrencamp, S.L. (2011). *Principles of animal communication*, 2nd ed  
632 (Sinauer Associates).
- 633 10. Weber, E.H. (1834). *De Pulsu, resorptione, auditu et tactu: Annotationes anatomicae et*  
634 *physiologicae ...* (C.F. Koehler).
- 635 11. Akre, K.L., and Johnsen, S. (2014). Psychophysics and the evolution of behavior. *Trends in*  
636 *Ecology & Evolution* 29, 291–300. 10.1016/j.tree.2014.03.007.
- 637 12. Wyttenbach, R.A., May, M.L., and Hoy, R.R. (1996). Categorical Perception of Sound Frequency  
638 by Crickets. *Science* 273, 1542–1544.
- 639 13. Weary, D.M. (1990). Categorization of song notes in great tits: which acoustic features are used  
640 and why? *Animal Behaviour* 39, 450–457. 10.1016/S0003-3472(05)80408-7.
- 641 14. May, B., Moody, D.B., and Stebbins, W.C. (1989). Categorical perception of conspecific  
642 communication sounds by Japanese macaques, *Macaca fuscata*. *J Acoust Soc Am* 85, 837–847.  
643 10.1121/1.397555.
- 644 15. Nelson, D.A., and Marler, P. (1989). Categorical perception of a natural stimulus continuum:  
645 birdsong. *Science* 244, 976–978. 10.1126/science.2727689.
- 646 16. Ehret, G. (1992). Categorical perception of mouse-pup ultrasounds in the temporal domain.  
647 *Animal Behaviour* 43, 409–416. 10.1016/S0003-3472(05)80101-0.
- 648 17. Fischer, J. (2006). Categorical Perception in Animals. *Encyclopedia of Language & Linguistics*.  
649 10.1016/B0-08-044854-2/00806-3.

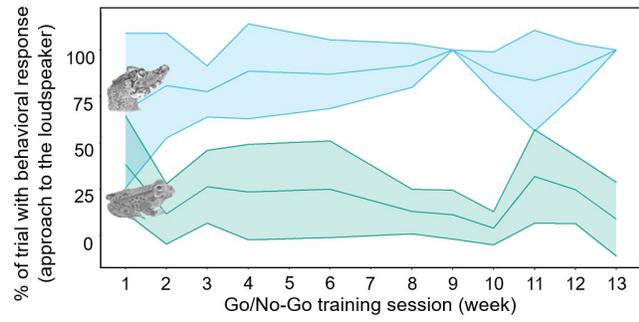
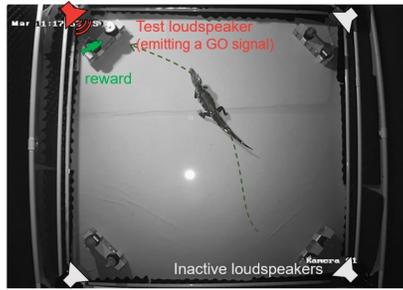
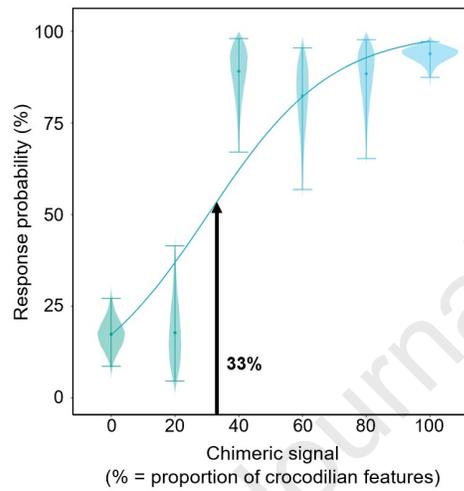
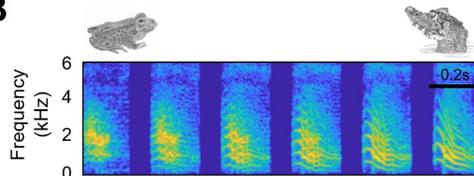
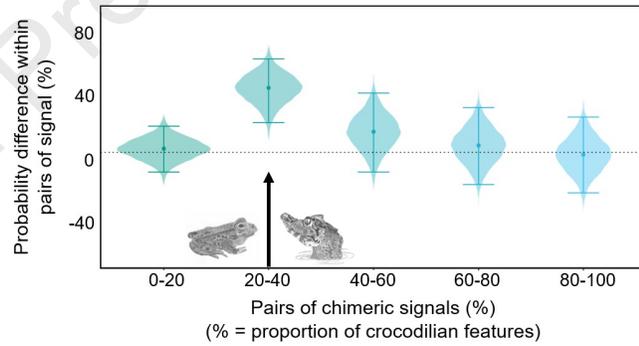
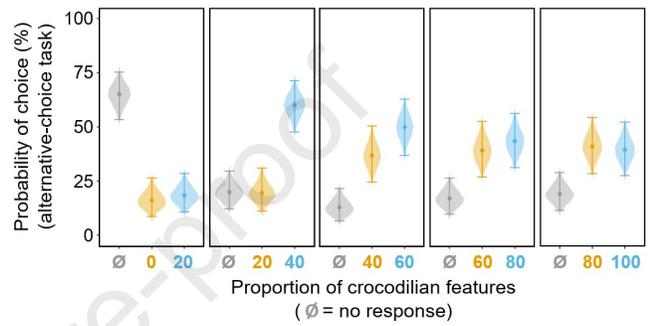
- 650 18. Baugh, A.T., Akre, K.L., and Ryan, M.J. (2008). Categorical perception of a natural, multivariate  
651 signal: mating call recognition in túngara frogs. *Proc Natl Acad Sci U S A* *105*, 8985–8988.  
652 10.1073/pnas.0802201105.
- 653 19. Green, P.A., Brandley, N.C., and Nowicki, S. (2020). Categorical perception in animal  
654 communication and decision-making. *Behav Ecol* *31*, 859–867. 10.1093/beheco/araa004.
- 655 20. Reber, S.A. (2020). Crocodilians Are Promising Intermediate Model Organisms for Comparative  
656 Perception Research. *Comparative Cognition & Behavior Reviews* *15*, 111–129.  
657 10.3819/CCBR.2020.150004.
- 658 21. Vergne, A.L., Pritz, M.B., and Mathevon, N. (2009). Acoustic communication in crocodilians: from  
659 behaviour to brain. *Biological Reviews* *84*, 391–411. 10.1111/j.1469-185X.2009.00079.x.
- 660 22. Grigg, G.C., and Kirshner, D. (2015). *Biology and evolution of crocodylians* (Comstock Publishing  
661 Associates a division of Cornell University Press).
- 662 23. Bierman, H.S., and Carr, C.E. (2015). Sound localization in the alligator. *Hearing Research* *329*,  
663 11–20. 10.1016/j.heares.2015.05.009.
- 664 24. Papet, L., Grimault, N., Boyer, N., and Mathevon, N. (2019). Influence of head morphology and  
665 natural postures on sound localization cues in crocodilians. *Royal Society Open Science* *6*,  
666 190423. 10.1098/rsos.190423.
- 667 25. Papet, L., Raymond, M., Boyer, N., Mathevon, N., and Grimault, N. (2020). Crocodiles use both  
668 interaural level differences and interaural time differences to locate a sound source. *The Journal*  
669 *of the Acoustical Society of America* *148*, EL307–EL313. 10.1121/10.0001979.
- 670 26. Thévenet, J., Papet, L., Campos, Z., Greenfield, M., Boyer, N., Grimault, N., and Mathevon, N.  
671 (2022). Spatial release from masking in crocodilians. *Commun Biol* *5*, 1–14. 10.1038/s42003-022-  
672 03799-7.
- 673 27. Vergne, A.L., Aubin, T., Martin, S., and Mathevon, N. (2012). Acoustic communication in  
674 crocodilians: information encoding and species specificity of juvenile calls. *Anim Cogn* *15*, 1095–  
675 1109. 10.1007/s10071-012-0533-7.
- 676 28. Vergne, A.L., Aubin, T., Taylor, P., and Mathevon, N. (2011). Acoustic signals of baby black  
677 caimans. *Zoology* *114*, 313–320. 10.1016/j.zool.2011.07.003.
- 678 29. DeGroot, M.H. (1970). *Optimal statistical decisions*. (McGraw-Hill).
- 679 30. Akre, K.L., and Johnsen, S. (2016). Communication Through a Window of Error: Proportional  
680 Processing and Signal Categorization. In *Psychological Mechanisms in Animal Communication*  
681 *Animal Signals and Communication.*, M. A. Bee and C. T. Miller, eds. (Springer International  
682 Publishing), pp. 137–167. 10.1007/978-3-319-48690-1\_6.
- 683 31. Morse, P.A., and Snowdon, C.T. (1975). An investigation of categorical speech discrimination by  
684 rhesus monkeys. *Perception & Psychophysics* *17*, 9–16. 10.3758/BF03203991.
- 685 32. Sinnott, J.M., Beecher, M.D., Moody, D.B., and Stebbins, W.C. (1976). Speech sound  
686 discrimination by monkeys and humans. *The Journal of the Acoustical Society of America* *60*,  
687 687–695. 10.1121/1.381140.

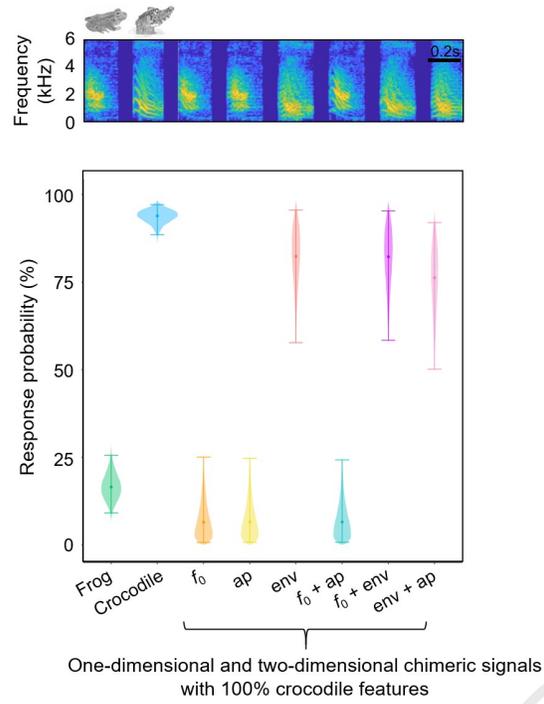
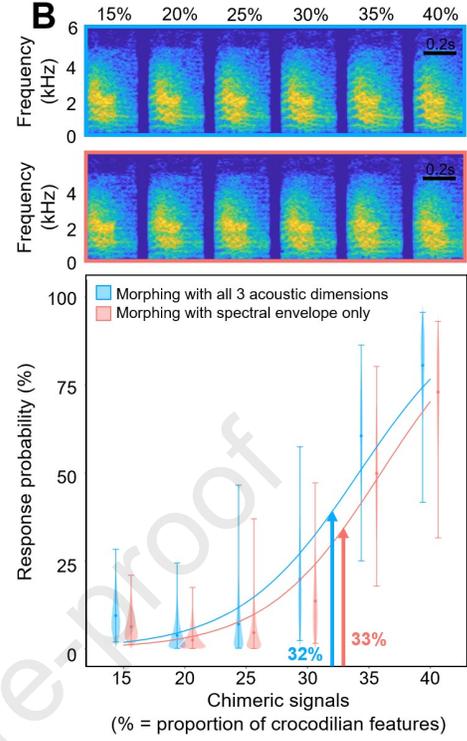
- 688 33. Kuhl, P.K., and Miller, J.D. (1978). Speech perception by the chinchilla: Identification functions for  
689 synthetic VOT stimuli. *The Journal of the Acoustical Society of America* 63, 905–917.  
690 10.1121/1.381770.
- 691 34. Kluender, K.R., Diehl, R.L., and Killeen, P.R. (1987). Japanese quail can learn phonetic categories.  
692 *Science* 237, 1195–1197. 10.1126/science.3629235.
- 693 35. Wetzell, W., Wagner, T., Ohl, F.W., and Scheich, H. (1998). Categorical discrimination of direction  
694 in frequency-modulated tones by Mongolian gerbils. *Behavioural Brain Research* 91, 29–39.  
695 10.1016/S0166-4328(97)00099-5.
- 696 36. Furuyama, T., Kobayasi, K.I., and Riquimaroux, H. (2017). Acoustic characteristics used by  
697 Japanese macaques for individual discrimination. *Journal of Experimental Biology* 220, 3571–  
698 3578. 10.1242/jeb.154765.
- 699 37. Burgering, M.A., Vroomen, J., and ten Cate, C. (2019). Zebra finches (*Taeniopygia guttata*) can  
700 categorize vowel-like sounds on both the fundamental frequency (“pitch”) and spectral  
701 envelope. *Journal of Comparative Psychology* 133, 106–117. 10.1037/com0000143.
- 702 38. Fischer, J. (1998). Barbary macaques categorize shrill barks into two call types. *Animal Behaviour*  
703 55, 799–807. 10.1006/anbe.1997.0663.
- 704 39. Fitch, W.T., and Giedd, J. (1999). Morphology and development of the human vocal tract: a study  
705 using magnetic resonance imaging. *J Acoust Soc Am* 106, 1511–1522. 10.1121/1.427148.
- 706 40. Smith, D.R.R., Patterson, R.D., Turner, R., Kawahara, H., and Irino, T. (2005). The processing and  
707 perception of size information in speech sounds. *J Acoust Soc Am* 117, 305–318.  
708 10.1121/1.1828637.
- 709 41. Chabert, T., Colin, A., Aubin, T., Shacks, V., Bourquin, S.L., Eley, R.M., Acosta, J.G., and  
710 Mathevon, N. (2015). Size does matter: crocodile mothers react more to the voice of smaller  
711 offspring. *Sci Rep* 5, 15547. 10.1038/srep15547.
- 712 42. Carr, C.E., and Christensen-Dalsgaard, J. (2015). Sound Localization Strategies in Three Predators.  
713 *BBE* 86, 17–27. 10.1159/000435946.
- 714 43. Thévenet, J., Kehy, M., Boyer, N., Pradeau, A., Papet, L., Gaudrain, E., Grimault, N., and  
715 Mathevon, N. (2023). Data from: Sound categorization by crocodilians [Data set]. Zenodo.  
716 <https://doi.org/10.5281/zenodo.7699715>.
- 717 44. Kawahara, H., Masuda-Katsuse, I., and de Cheveigné, A. (1999). Restructuring speech  
718 representations using a pitch-adaptive time–frequency smoothing and an instantaneous-  
719 frequency-based F0 extraction: Possible role of a repetitive structure in sounds1Speech files  
720 available. See <http://www.elsevier.nl/locate/specom1>. *Speech Communication* 27, 187–207.  
721 10.1016/S0167-6393(98)00085-5.
- 722 45. Kawahara, H. (2006). STRAIGHT, exploitation of the other aspect of VOCODER: Perceptually  
723 isomorphic decomposition of speech sounds. *Acoustical Science and Technology* 27, 349–353.  
724 10.1250/ast.27.349.
- 725 46. Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of*  
726 *Statistical Software* 80, 1–28. 10.18637/jss.v080.i01.
- 727 47. Kruschke, J. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan* (Academic  
728 Press).

- 729 48. Kruschke, J.K., and Liddell, T.M. (2018). The Bayesian New Statistics: Hypothesis testing,  
730 estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic*  
731 *Bulletin & Review* 25, 178–206. 10.3758/s13423-016-1221-4.
- 732 49. Puig-Diví, A., Escalona-Marfil, C., Padullés-Riu, J.M., Busquets, A., Padullés-Chando, X., and  
733 Marcos-Ruiz, D. (2019). Validity and reliability of the Kinovea program in obtaining angles and  
734 distances using coordinates in 4 perspectives. *PLOS ONE* 14, e0216448.  
735 10.1371/journal.pone.0216448.
- 736

Journal Pre-proof



**A****B****C**

**A****B**

## HIGHLIGHTS

- Crocodylians can categorize sounds along an acoustic continuum.
- This categorization is solely based on the spectral envelope of sound signals.
- The boundary between two sound categories may be influenced by learning.

Journal Pre-proof