



**HAL**  
open science

# Auditory perception and decision-making in crocodilians

Julie Thévenet

► **To cite this version:**

Julie Thévenet. Auditory perception and decision-making in crocodilians. *Neurons and Cognition* [q-bio.NC]. Université Claude Bernard - Lyon I, 2022. English. NNT : 2022LYO10091 . tel-04166018

**HAL Id: tel-04166018**

**<https://theses.hal.science/tel-04166018>**

Submitted on 19 Jul 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.



# THESE de DOCTORAT DE L'UNIVERSITE CLAUDE BERNARD LYON 1

**Ecole Doctorale 162  
Mécanique, Energétique, Génie civil, Acoustique**

**Discipline** : Bioacoustique

Soutenue publiquement le 09/11/2022, par :

**Julie Thévenet**

---

## **Perception auditive et prise de décision chez les crocodiliens**

*Auditory perception and decision-making in crocodilians*

---

Devant le jury composé de :

**Jean-Christophe BERA**

Professeur des universités, Université Lyon 1

**Président**

**Sophie LUMINEAU**

Maîtresse de conférences, Université de Rennes 1

**Rapporteuse**

**Marco GAMBA**

Professeur associé, Université de Turin

**Rapporteur**

**Fanny RYBAK**

Maîtresse de conférences, Université Paris-Saclay

**Examinatrice**

**Nicolas GRIMAUULT**

Directeur de recherche, CNRS, CRNL Lyon

**Directeur de thèse**

**Nicolas MATHEVON**

Professeur des universités, Université de Saint-Etienne

**Directeur de thèse**



# Remerciements

Quoi de plus cocasse que de faire une thèse sur la communication quand soi-même, on a du mal à communiquer... Je crois bien avoir tout de même un peu progressée tout au long de ces trois dernières années (ou pas trop, vous me direz). La thèse, c'est dur, mais ça fait grandir. A tous points de vue.

Je tiens à remercier avant tout les membres de mon jury de thèse : *Sophie Lumineau* et *Marco Gamba* d'avoir accepté d'évaluer cette thèse, ainsi que *Fanny Rybak* et *Jean-Christophe Bera* de faire partie de mon jury. J'espère que vous aurez plaisir à lire ce manuscrit.

Un énorme merci à mes directeurs de thèse, *Nicolas Grimault* et *Nicolas Mathevon*. Merci de m'avoir ouvert les portes de vos univers respectifs, de m'avoir donné cette chance inouïe de travailler sur un tel sujet, de m'avoir fait confiance tout au long de ces 3 ans et de m'avoir permis d'apprendre tant de choses et de vivre des aventures si exceptionnelles. Merci aussi au labex CeLyA d'avoir financé ce projet qui sort des cadres.

Tout avait commencé par un simple email. J'avais décidé de contacter le laboratoire ENES. Très surprise de la réponse quasi-immédiate de Nicolas Mathevon, nous fixons une date pour que je puisse venir au laboratoire afin de discuter plus en détails. Quel n'est pas mon étonnement quand je vois un grand barbu, à l'allure détente, mais un je ne sais quoi qui impressionne quand même (la barbe sûrement), qui me dit que oui, il aura besoin d'un stagiaire de M2 pour l'aider à travailler sur ses crocos. Des crocos ?! C'est fouuuuuuu ! C'était parti pour mon stage de Master avec l'ENES. *Les Nicolas* et *Léo* m'avaient ouvert les portes d'une grande aventure scientifique et humaine. Léo m'a donné une confiance aveugle tout le long de mon stage, m'a aidé, supporté (dans tous les sens du terme), m'a transmis ses connaissances tant scientifiques qu'en musique, rando, jardinage et j'en passe. Un énorme merci à toi pour tout ça et d'avoir largement pesé dans la balance en ma faveur lorsque l'idée de la thèse fut évoquée par la team des Nicolas. Tout ça, c'est un peu grâce à toi quand même.

Un dernier Nicolas vient s'ajouter à la team. *Neeko*, alias Iguanamousse (avec l'accent qui fait bien). Sans toi, rien n'aurait été possible. Merci d'avoir été mon pilier crocodilo-technico-émotionnel tout au long de ces 3 années et de m'avoir tant transmis. Nos fous rires, nos épopées



marocaines et nos pics d'adrénaline à la capture des petits monstres vont me manquer. Ne change jamais. Gloire à Skipi.

En parlant de manipulation de crocos, un grand merci *Aurélie* merci d'avoir été là en soutien. Merci pour ce que tu as pu me transmettre et d'avoir partagé toutes tes histoires plus loufoques les unes que les autres.

Une partie de ce manuscrit n'aurait pas pu voir le jour sans *Mounia*. Merci d'avoir été une stagiaire en or, autonome, hyper efficace et sans peur. Ma dernière manip en labo n'aurait jamais été aussi bien conduite que par toi, ton aide a été considérable.

Un grand merci à toute l'équipe de l'ENES et de CAP, pour ces 3 années exceptionnelles passées en votre compagnie, pour le partage tant scientifique qu'humain (et culinaire) et pour votre bonne humeur au quotidien ! J'ai découvert des personnes toutes plus extraordinaires les unes que les autres, des stagiaires, aux doctorants et post-doctorants, *Anna, Arthur, Clément, Joël* (oui, tu es encore jeune), *Loïc, Lucie, Mathilde, Siloé, Théophile*, du côté des Stéphanois et *Anna, Aurélien, Cyriaque, Florian, Margot* du côté des Lyonnais. *Emi*, que te dire. Cette dernière année a été difficile pour toutes les deux, mais on en sort que plus fortes que jamais et on ne peut être que fières de tout ce qu'on a accompli. Merci d'avoir été mon premier soutien avant même le début de cette thèse.

Enfin, un immense merci à tous nos collaborateurs et aux personnes qui ont rendu tous ces travaux possibles : *Olivier Marquis* et le Parc Zoologique de Paris pour nous avoir fournis les 6 caïmans nains avec lesquels j'ai eu l'immense chance de travailler à l'ENES. Ils se sont avérés être des collaborateurs en or pour cette thèse. Toute l'équipe du Crocoparc, pour nous avoir permis de travailler dans des conditions royales sur leur plus que nombreux Crocodiles du Nil. Merci pour leur accueil toujours aussi chaleureux et leur support technique et logistique à toute épreuve. *Paulo Fonseca, Pedro Fonseca* et *Ernesto Tualufo* pour ce séjour au Mozambique, riche scientifiquement, culturellement et humainement. Un énorme merci d'avoir permis ce voyage et d'avoir rendu les choses si faciles sur place.

À présent, à situation un peu particulière, remerciements un peu particuliers... Je ne vous ai pas oubliés, pas de panique !

Mon soutien inconditionnel de fin de thèse, *Naïs*, tu as ma reconnaissance éternelle pour tout ce que tu as fait pour moi en ces derniers mois un peu compliqués... Ce fut une rencontre qui est allée de surprises en surprises. On n'avait pourtant pas grand-chose en commun aux premiers abords mis à part cette obsession pour les chats noirs. Puis de terrain marocain en colocation, c'en est devenu flippant de voir à quel point on pouvait se ressembler parfois (souvent ?). Ton aide a été et reste inestimable à tout point de vue, MERCI.

*Lény*, mon autre pilier de fin de thèse et pas des moindres. Merci pour absolument tout ce que tu as fait pour moi, de m'avoir prêté une oreille attentive et sans jugement, de m'avoir tant posé de questions toujours autant dans le mille, et de m'avoir soutenue et supportée alors que tout ce que je voulais, c'était danser dans ma chambre. Merci d'avoir été là en toutes circonstances.

*Floriane*, nos chemins ont failli se croiser un peu plus tôt, mais ce contretemps n'a été que l'occasion de mieux se découvrir en cette dernière année un peu spéciale. Merci pour ton soutien, pour tous ces moments partagés, les très bons comme nos moments randos, bords de Loire, chips (et sandwich de chips), cochonou, montage de bibliothèque, de pizzas sans four, tout comme les moins bons quand plus rien n'allait. Une mention spéciale à tous nos moments beaux, de loin les meilleurs.

*Coralynn*, un gigantesque merci... Ce fameux petit séjour chez vous, nos conversations et nos sessions jardinage ont contribué à mon déclic personnel. Il n'y a pas d'autres mots que merci à toi et *Léo*, vous m'avez inspirée, accompagnée et soutenue à tout point de vue. Tout simplement, je crois bien que vous êtes les meilleurs. Merci...

Comment ne pas te remercier aussi *Léo*, pour nos sessions de bloc, de débriefings, de messages à 3h du matin, de Beaufort, de bouées, de randos épiques, et j'en passe... Merci de m'avoir proposé ton aide à tout moment, d'avoir toujours été là, et de m'avoir fait découvrir tout pleins de belles personnes.

*Sans vous, ce manuscrit n'aurait sûrement pas pu voir le jour... Je vous remercie du fond du cœur.*



Un dernier remerciement à toute ma famille et particulièrement mes parents, qui m'ont toujours soutenu dans toutes les épreuves, qui m'ont apporté leur confiance dans les différentes phases de ces dernières années, alors même que je pouvais en manquer parfois (ou très souvent). *Kevin* et *Corentin*, merci de m'avoir supporté moi, ma folie, et tous mes rebondissements (au propre comme au figuré) tout ce temps !

Je suis à présent fière de vous présenter cette thèse, le travail de 3 années qui m'aura apporté certes des compétences scientifiques, mais aussi, indéniablement un peu plus d'assurance dans cette vie.

Ma petite mamie, je sais qu'on se ressemble beaucoup, mais sois en rassurée, cette thèse aura eu le mérite de m'apporter confiance en moi et moins de doutes pour aller à fond dans cette vie.

# Résumé

Nous vivons, comme tout le règne animal, dans des environnements complexes formés de nombreuses informations multisensorielles. Parmi ces informations, les crocodiliens sont particulièrement aptes à utiliser les signaux acoustiques afin de sonder leur environnement. L'évolution a en effet doté ces animaux de bonnes capacités auditives, leur accordant d'utiliser la communication acoustique activement au cours de leur vie. La biomécanique associée à leur production vocale n'est pas encore très bien comprise, mais quoi qu'il en soit, les crocodiliens sont capables d'émettre différents types de vocalisations ayant chacun un rôle fonctionnel bien distinct. Les jeunes crocodiliens en particulier utilisent diligemment un panel de vocalisations qui sont essentielles à la survie. Ils émettent notamment des cris de contacts qui assurent leur cohésion et des cris de détresse leur permettant de signaler un danger et d'appeler une aide parentale. Du point de vue du récepteur, l'anatomie de l'oreille moyenne et interne des crocodiliens est complexe et leur permet d'avoir une large sensibilité auditive (0.1-3 kHz, avec un maximum de sensibilité autour de 1 kHz correspondant à la bande de fréquence dominante des cris de détresse). De manière plus poussée, des études ont montrés qu'ils possèdent également des adaptations leur permettant de localiser efficacement et avec une bonne précision les sources sonores. Ces capacités constituent un réel avantage pour l'analyse de leur environnement sonore puisqu'ils peuvent ainsi potentiellement ségréger des sons provenant de sources spatialement distinctes et localiser ainsi seulement les sources de sons qui sont biologiquement pertinentes.

Cependant, les crocodiliens font face à un afflux massif d'information auditive, qu'ils doivent détecter, identifier puis trier. Les sons qui sont d'intérêt pour eux peuvent être masqués par des signaux non pertinents. Ce bruit de fond environnant peut donc grandement affecter la communication acoustique, et peut être critique notamment pour les plus jeunes pour qui se tromper dans la détection, la localisation ou l'identification d'un signal peut être fatal. Des signaux sonores peuvent être extraits du bruit environnant de par leur localisation (deux signaux spatialement distincts ou un signal qui se trouve éloigné de la source du bruit peuvent être plus facilement distingué : principe du démasquage spatial), mais aussi de par leurs caractéristiques acoustiques. Cependant, on ne sait toujours pas comment les crocodiliens prennent des décisions quant à quels signaux sont d'intérêt et lesquels ne le sont pas. Comment est-ce qu'ils décodent l'information et sur quels indices acoustiques est-ce qu'ils se basent pour identifier et classer un signal ? Enfin, on ne sait également pas s'ils sont aussi capables d'utiliser des





informations qui ne leur sont *a priori* pas destinées, comme des vocalisations produites par d'autres animaux, ce qui contribuerait à leur position de top-prédateur.

Afin de mieux comprendre les performances de détection de signaux dans le bruit, d'identification et de tri de l'information acoustique, nous avons mené une série de trois études. Tout d'abord, nous avons mené des expériences de détection de signaux acoustiques en environnement bruyant. Cette première étude a été réalisée en combinant trois approches : une expérience de playback en nature, où des femelles yacare caïmans (*Caiman yacare*) devaient répondre à des cris de détresse masqués par un bruit continu ; une seconde expérience de playback en zoo, où des crocodiles du Nil (*Crocodylus niloticus*) juvéniles devaient répondre à des cris de contact masqués par un bruit ; et une dernière expérience en laboratoire, où 2 crocodiles du Nil ont été entraînés à répondre à des signaux synthétiques. Dans tous les contextes, la séparation spatiale de la source du signal par rapport à celle du bruit améliorerait les performances de détection, confirmant le mécanisme de démasquage spatial qui leur permettrait d'améliorer l'analyse de la scène auditive.

Ensuite, je me suis intéressée au codage de l'information et de la manière dont de jeunes crocodiliens peuvent identifier un cri de conspécifique. Un des mécanismes cognitifs qui permet d'identifier et de réagir rapidement aux informations contenues dans un flux de signaux est la catégorisation, où l'individu récepteur perçoit un continuum de signaux comme appartenant à des catégories discrètes. Au travers d'une expérience de playback en zoo avec de jeunes crocodiles du Nil, ainsi que de plusieurs expériences en laboratoire impliquant un conditionnement de jeunes caïmans nains, j'ai donc testé si les crocodiliens sont capables de partitionner un continuum acoustique en catégories significatives et s'ils utilisent un paramètre acoustique en particulier pour baser leur classification. Les résultats montrent que les jeunes crocodiles confrontés à un continuum acoustiques chimérique allant d'un cri de grenouille à un cri de crocodile classent chaque variante acoustique dans l'une de ces deux catégories en se basant uniquement sur leur enveloppe spectrale, établissant ainsi une frontière nette là où aucune frontière acoustique n'existe. La frontière perceptuelle entre les deux catégories n'est cependant pas la même entre les expériences de playback et de go/no-go en laboratoire, suggérant que le contexte, voire l'apprentissage peuvent moduler les frontières entre des catégories de sons et jouer ainsi un rôle dans la prise de décision rapide.

Enfin, la dernière étude vise à explorer la perception de l'environnement sonore des crocodiliens dans un champ plus large. Grâce à des expériences de playback, j'étudie leur

capacité à utiliser des informations véhiculées par des sons produits par d'autres animaux qui peuvent être présents dans leur environnement. Les crocodiliens adultes étant fortement réceptifs aux cris de détresse des juvéniles, je me suis focalisée sur leur perception de la détresse encodée dans les vocalisations d'espèces phylogénétiquement très distantes. Les résultats montrent que les crocodiles sont attirés par des cris de bébés hominidés (bonobo, chimpanzé et humain), et que l'intensité de leur réponse dépend d'un set de paramètres acoustiques spécifiques (principalement caractérisant la rugosité des cris), qui sont probablement des marqueurs universels de la détresse. Les crocodiles sont donc sensibles au degré de détresse encodé dans les vocalisations de vertébrés phylogénétiquement très éloignés, démontrant leurs capacités à tirer profit des informations acoustiques présentes dans leur environnement.

Pour conclure, toutes ces études contribuent à mieux comprendre comment les crocodiliens peuvent percevoir leur environnement sonore et démontrent qu'ils peuvent utiliser les informations encodées dans des signaux acoustiques afin de prendre des décisions rapides. Tout au long de cette thèse, le facteur contextuel a eu un grand impact sur les résultats obtenus, suggérant qu'on ne peut dissocier les processus de perception de l'individualité et du vécu de l'individu.





# Contents

<b>I. General introduction .....</b>	<b>11</b>
1. Overview of the acoustic world of crocodilians.....	11
1.1 Phylogeny and terminology .....	11
1.2 Auditory and vocal system.....	13
1.3 Vocal repertoire.....	16
2. Perception and response to sounds by crocodilians.....	20
2.1 Sound signals detection.....	20
2.2 Conspecifics information identification and decision-making.....	23
2.3 Taking advantage of the sound environment: perceive the others.....	25
<b>II. General methods .....</b>	<b>27</b>
1. Playback experiments .....	27
2. Laboratory experiments .....	28
3. Timeline of the experiments.....	35
<b>III. Sound detection in noise.....</b>	<b>37</b>
1. Introduction.....	39
2. Results .....	42
2.1 Spatial Release from Masking by adult crocodilians .....	42
2.2 Spatial Release from Masking by juvenile crocodilians .....	46
2.3 Spatial Release from Masking by crocodilians to detect a non-biological signal..	48
3. Discussion.....	51
4. Methods .....	55
5. Supplementary information.....	66
<b>IV. Sound categorization .....</b>	<b>73</b>
1. Ethological experiments.....	73
1.1 Introduction.....	75



1.2	Methods .....	77
1.3	Results .....	86
1.4	Discussion.....	90
1.5	Supplementary information.....	94
2.	Pupillometry experiments .....	96
2.1	Pupillometry principle .....	96
2.2	Particularity of the crocodilian’s eye .....	97
2.3	Methodology .....	98
2.4	Preliminary results.....	101
2.5	Perspectives .....	103
<b>V.</b>	<b>Hearing others: perceiving heterospecific distress.....</b>	<b>105</b>
1.	Introduction.....	107
2.	Methods .....	110
3.	Results .....	113
4.	Discussion.....	116
5.	Supplementary information.....	120
<b>VI.</b>	<b>General discussion.....</b>	<b>125</b>
	<b>References.....</b>	<b>131</b>
	<b>Annex.....</b>	<b>147</b>

# I. General introduction

In our daily environment, we have to decide whether a signal that occurs is relevant to us or not. All behavior and actions are based on more or less considered decisions, whether it is choosing the name of one's crocodile, having the courage to catch one, or a more committed choice, such as whether or not to engage in a PhD. In the animal kingdom, animals also face situations in which they must make decisions. Discerning the decision-making processes underway or, even more so, being able to predict choices are fundamental to understanding cognition and behavior. While more is known about crocodilian language to date, we still do not know how they perceive their sound environment, nor how they use the information encoded in sound signals to make decisions. It seemed to me important and interesting during my thesis to try to give an account of what crocodilians could perceive of their acoustic world, how they detect the signals present in their environment, but also how they manage to sort the auditory information to keep only what is relevant for them. This thesis has an undeniable acoustic tint, but the behavioral aspect is also present and was close to my heart.

## 1. Overview of the acoustic world of crocodilians

### 1.1 Phylogeny and terminology

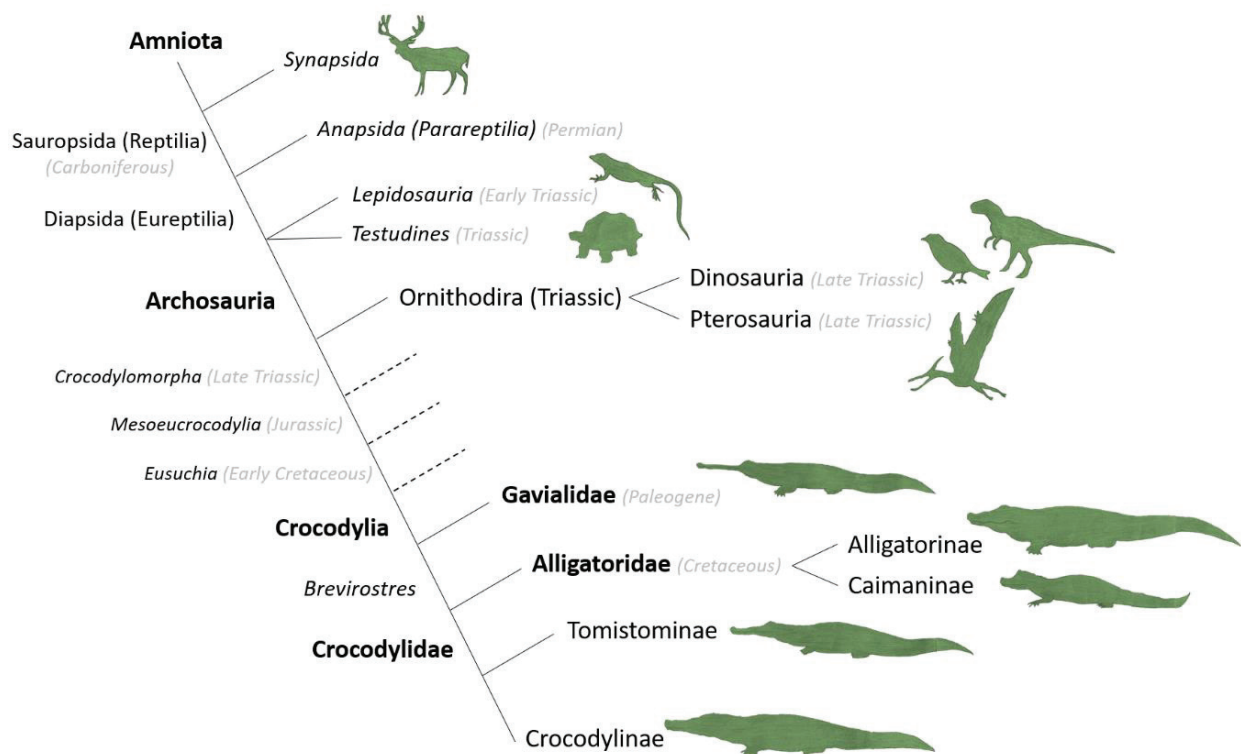
First of all, before talking about crocodiles and caimans, I have to introduce you a little bit this so interesting taxonomic group which occupied me these last 3 and a half years of my life. Just a terminology point, crocodilians, as we will call them throughout this thesis, is a familiar term to refer to the living species belonging the order of Crocodylia (Figure I.1; Stevenson, 2019).

245 million years ago, or about 185 million years before the largest mass extinction event on Earth, crocodylians shared a common ancestor with birds and dinosaurs, defining the archosaur group (Hugall et al., 2007; Nesbitt, 2011). Approximately 65 million years ago, at the end of the Cretaceous period, a large asteroid hit the Earth, causing a dramatic deterioration in atmospheric and oceanic conditions and a drastic change in climate. This impact caused the disappearance of most of the Archosaurs, with the exception of birds and a small number of crocodile species that could have avoided the worst of the climatic consequences. When



conditions became favourable enough again, the remaining crocodylian species were able to adapt to the new conditions and prosper. Crocodylians then diverged into three families that survived into modern times: the Alligatoridae (including alligators and caimans), the Crocodylidae (the "true" crocodiles), and the Gavialidae (the gharials).

The Gavialidae diverged first about 70 million years ago, while the Crocodylidae and Alligatoridae separated about 25 million years later (Grigg & Kirshner, 2015). The timing of the divergence of these three families is still unclear and there is no direct chain of descent between ancient and modern crocodylian species. To date, crocodylians remain the closest living reptilian relatives to dinosaurs (Hugall et al., 2007).



**Figure I.1.** Simplified cladogram of Archosauria (based on Grigg & Kirshner, 2015; Stevenson, 2019).

For now, and pending further discoveries resulting from ever more advanced combined methods of genetic, morphological and biogeographic analysis, 28 species of crocodylians are recognized, including 19 species of Crocodylidae, 8 species of Alligatoridae (with possibly four subspecies of the species *Caiman crocodylus*) and one species of Gavialidae (Stevenson, 2019; Murray et al., 2019; Reber, 2020). Yet, the number of fossil species is incredibly higher (tenfold), compared to the number of extant crocodylians (Stevenson, 2019).

The three families of extant crocodylians have similar general biology and behavior. They are semiaquatic quadrupedal apex predators that live at the interface between air and water (Grigg & Kirshner, 2015). While adults crocodylians have only few predators (mostly humans or other crocodylians), eggs, hatchlings or juveniles are more prone to predations and killed by a variety of invertebrates, fishes, amphibians, birds, reptiles or mammals (Somaweera et al., 2013).

Maybe because they are ambush predators, when they are in the water, their body is almost entirely submerged, with only their eyes, nostrils, and ears just above the surface of the water: all their senses are on alert. For now, if their sensory organs are better and better studied from an anatomical and evolution point of view, their perceptual capacities, in all senses, remain poorly known. Yet, the difficulty of working with such huge and dangerous animals makes behavioral experiments long and complicated to set up. Although the sensory world of crocodylians is incredibly diverse, I will focus on their auditory perception, from a brief description of the anatomy and physiology of their vocal and auditory organs, to the behavioral experiments that have already been conducted on that topic.

## **1.2 Auditory and vocal system**

### **1.2.1 Physiology of the ear**

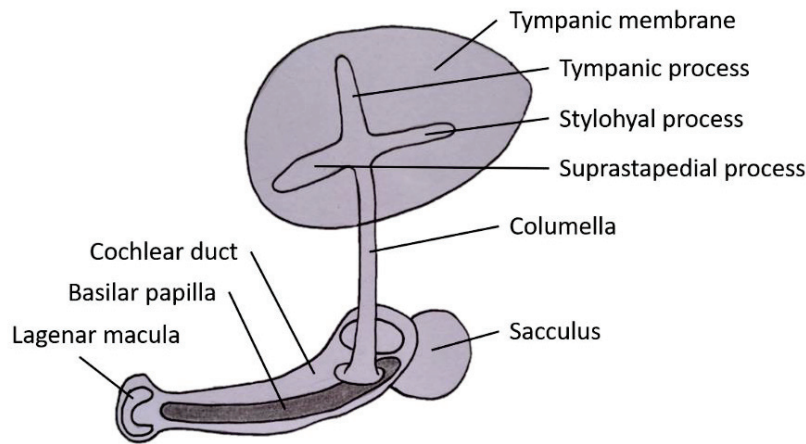
The external ears of crocodylians are located just behind the eyes, on the top of the head. They have this particularity to be covered by muscular flaps (also called “ear lids”; Saunders et al., 2000). The flaps close when the animal submerges, and relax when out of the water, leaving only a small horizontal opening (Saunders et al., 2000; Wever & Vernon, 1957). This control of the slit opening protects the direct underlying tympanic membrane and prevents water entry when the animal dives.

The middle and inner ear is anatomically quite similar to that of birds, suggesting that this apparatus is ancestral to all archosaurs, including extinct ones (Figure I.2; Wever & Vernon, 1957; Baird, 1974; Gleich & Manley, 2000; Saunders et al., 2000). The middle ear of crocodylians consists of a single ossicle, the columella, corresponding to the stapes of the mammalian ear (the smallest and innermost of the bones in the chain of three ossicles of the mammalian ear; Saunders et al., 2000). The cartilaginous extracolumella connects a bony columellar shaft with three anterior processes: the tympanic, stylohyal and extrastapedial processes. An extra-columellar muscle is connected to the suprastapedial process, and would





attenuate pressure waves on the tympanic membrane, as does the stapedial reflex in humans (Saunders et al., 2000). The columellar shaft terminates posteriorly in large foot plate.



**Figure I.2.** Schematic representation of the auditory apparatus in crocodilian: middle and inner ear (adapted from A. Vergne et al., 2009; Grigg & Kirshner, 2015).

The inner ear consists on a curved cochlea as in birds and other reptiles, instead of a coiled one as in mammals. The hair cells cover a basilar papillae (A. Vergne et al., 2009) and are of two anatomic types: short hair cells, found predominantly over the basilar papilla, and tall hair cells, less specialized and found mainly at the apex (Baird, 1974).

As in birds, the basilar papilla is organized topographically: the base, which has more short hair cells, is more sensitive to high frequencies than the apex with its many long hair cells. The large surface area of the basilar papilla allowed by the evolution of an elongated cochlea serves a wide range of hearing and good hearing sensitivity (see *part I.2.1*).

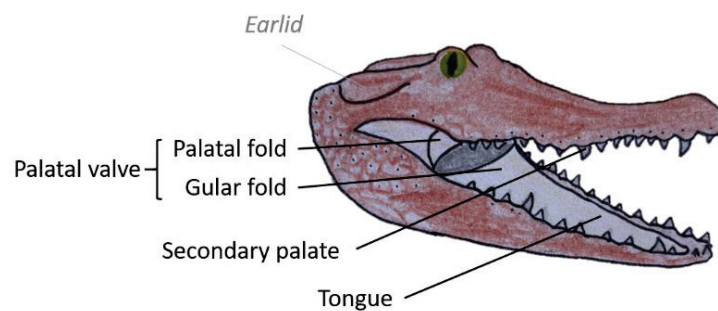
Two vestibular organs whose role in sound perception is unclear, are connected to the inner ear of crocodilians: the lagenar macula (at the apical pole) and the sacculus (at the base). The sacculus might be involved in the perception of low frequency sounds, including infrasound (see *part I.1.3* on adult long-distance call; Gleich & Manley, 2000; Todd, 2007).

Finally, the neurological pathway from the auditory nerve to the central auditory system is very similar of that of birds, with a tonotopic organization of the cochlear nuclei (see A. Vergne et al., 2009 for a review).

## 1.2.2 Physiology of the vocal production

The vocal apparatus of crocodilians evolved in a different way from that of birds. Crocodilians do not possess a dedicated vocal organ for sound production, producing their vocalizations through a larynx homologous to that of mammals (Reese, 1945). The vocal production is thought to involve airflow through the glottis, which is lined with membranous folds, and the contraction of muscles that varies the tension and modify the acoustic properties of the calls (Naifeh et al., 1970).

Looking in more detail at the anatomy of crocodilians, at the back of their oral cavity, the tongue is connected to what is known as the “palatal valve” (also called the gular valve), which corresponds to a basihyoid cartilage framed by epithelium and connective tissue (Figure I.3; Reese, 1945; Britton, 2001).



**Figure I.3.** Schematic head of a crocodile showing the structure associated with vocal production (based on Reese, 1945; Vergne et al., 2009; Grigg & Kirshner, 2015).

The palatal valve can be lifted against the posterior end of the secondary palate, called the “palatal fold” (which is also an extension of the connective tissue). The position of the palatal valve against the palate and thus its openness allows crocodilians to seal their pharyngeal cavity and prevents water from entering the mouth when crocodiles dive with their mouth open. This mechanism could also probably play a role in the modulation of vocal production, but for sure, the whole mechanics of vocal production in crocodilians remains still to be demonstrated.



## 1.3 Vocal repertoire

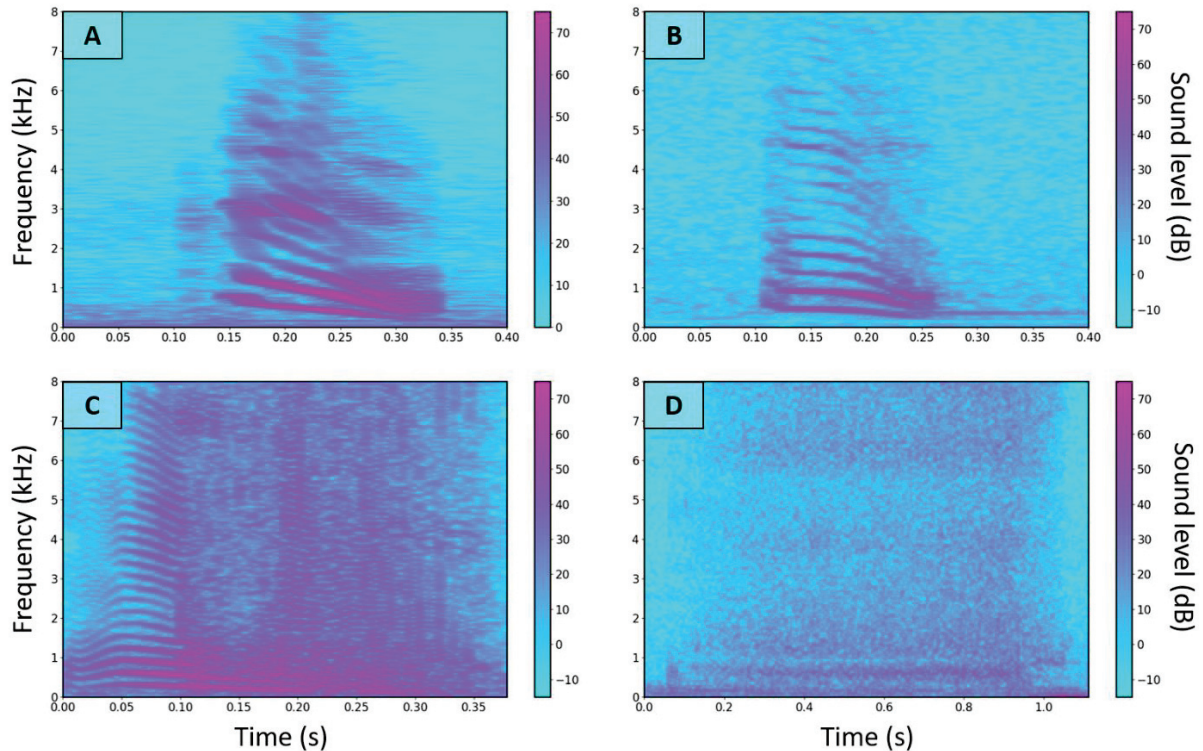
Reptiles are considered to be among the most silent vertebrates (Russell & Bauer, 2021) Their vocalizations are most often associated with basic behavior such as defensive/aggressive behaviors or mating activities, nonetheless, some lizards and more importantly crocodilians are the exception. All Crocodilians, all 28 species in the order, seem to emit vocalizations. Crocodilians are not only the most vocal reptiles. Somewhat surprisingly when one is unfamiliar with these animals, they possess a true repertoire of calls, with distinct functional roles and which appears to be universal among all crocodilians (Vergne et al., 2012). Acoustic communication is intense and has a major role in the early life of juveniles, when individuals are subject to heavy predation and must remain as cryptic as possible (Somaweera et al., 2013). As adults, crocodilians vocalize less than in their first years (*Crocodylus* species even less than Alligatoridae) and their repertoire evolves. To date, however, the vocal behavior of adults has been poorly studied and research has focused mainly on the repertoire of juveniles. I will present below the types of calls currently identified in the crocodilian repertoire. However, the classification of the calls into distinguishable acoustic and functional categories is not straightforward, and many of the call types identified by observations still require advanced acoustic analysis and experimental studies.

### 1.3.1 Juveniles acoustic communication

While the repertoire of juvenile crocodiles is varied, it is still limited to four main types of calls. All calls shared the same basic acoustic structure: a complex sound (a fundamental frequency with several harmonics) modulated in frequency and covering a wide range of frequencies (Figure I.4; Campbell, 1973; Britton, 2001; Vergne et al., 2012).

Let's start at the beginning. Crocodilians begin vocalizing very early in their life. In fact, they start vocalizing as early as inside the egg, up to five days before hatching (D. S. Lee, 1968; Somaweera & Shine, 2012). Female crocodilians lay their eggs in holes in sand or under building nests until the juveniles hatch (30-50 eggs on average). After 60 to 115 days of incubation (Stevenson, 2019), the eggs are ready to hatch. Just prior to doing so, the young begin to emit these **pre-hatching calls** (Figure I.4A) that are heard by nearby siblings within the others eggs, who begin to vocalize in turn. This emulation serve to synchronize hatchlings in a clutch but also warn the adult female of the imminence of hatching (Lee, 1968; Vergne & Mathevon, 2008). Playback experiments have previously shown both effects of these pre-

hatching calls: young were responsive to nearby hatching calls broadcast from a loudspeaker, as were adult females who approached and dug for the speaker hidden at the nest site (Vergne & Mathevon, 2008).



**Figure I.4.** Spectrograms of the main vocalizations of juvenile crocodilians (FFT window length of 1024 with 50% overlap). All calls shared the same basic acoustic structure: a fundamental frequency with several harmonics, modulated in frequency and covering a wide range of frequencies. **A.** Hatching call of a Nile crocodile *Crocodylus niloticus*. **B.** Contact call of a Nile crocodile. **C.** Distress call of a smooth-fronted caiman *Paleosuchus trigonatus*. The downward frequency modulation is preceded by a brief upward frequency modulation. **D.** Threat call of a smooth-fronted caiman.

Just after hatching, the hatchlings continue to call (**post-hatching calls**) to stimulate adult females, who come (or males in rare cases; see Lang et al., 1986) to carry the hatchlings in their mouth to the water. In addition to fine tune the hatching with the opening of the nest, hatching calls are also used by parents to distinguish infertile eggs and hence discarding them (Somaweera & Shine, 2012). Post-hatching calls are already acoustically different from pre-hatching calls (they are higher-pitched, longer in duration and with a higher intensity; Britton, 2001), but this is likely due to the egg shell that filters the calls rather than having a functional role. Nevertheless, it could potentially also be a cue for the siblings or the parent (Vergne et al.,



2009). With playback experiments, Vergne et al. (2007) showed that the identity of the juvenile is not encoded in these hatching calls. This demonstrated the interspecificity of the calls at the level of perception of juveniles and adult females.

Hatchlings and juveniles later emit **contact calls** (grunts or barks in older literature; Figure I.4B), high-pitched calls with a fairly low amplitude (Campbell, 1973; Herzog & Burghardt, 1977). This type of call can be emitted spontaneously several times per minute by juveniles when feeding, approaching other pod members, or moving in a group, and also elicits the approach of others. Thus, they might serve to ensure group cohesion (Campbell, 1973; Herzog & Burghardt, 1977; Britton, 2001).

Another type of call can be emitted by young crocodilians: the **distress calls** (Figure I.4C), which are louder than contact calls, with a higher starting frequency, a steeper frequency slope and a higher energy spectrum (Britton, 2001; Vergne et al., 2009). Downward frequency modulation can be preceded by an upward frequency modulation in some species such as the Nile crocodile *Crocodylus niloticus* (Campbell, 1973; Herzog & Burghardt, 1977; Vergne et al., 2007) and the Australian freshwater crocodile *Crocodylus johnstoni* (Britton, 2001). Furthermore, the intensity and pitch of calls can be modulated by mouth opening. Actually, the acoustic boundary between hatching calls, contact calls and distress calls is unclear. It is likely that the calls share a common basic acoustic structure and grade into each other (Campbell, 1973; Herzog & Burghardt, 1977; Vergne et al., 2012). Nonetheless, they are well perceived as distinct calls and elicit very different responses in juveniles (Vergne et al., 2011). Distress calls are produced in the presence of danger (e.g. predator threat, human capture). They have the demonstrated effect of warning other juveniles in a pod: juveniles stopped moving and began to emit distress calls in return when distress calls were played back (Vergne et al., 2011; Boucher et al., 2020). Unlike contact calls, females are sensitive to distress calls. Playbacks of distress calls caused the females to strongly approach the loudspeaker, suggesting that these calls induce parental protection (Vergne et al., 2011).

Finally, juveniles also emit **threat calls** or hissing (Figure I.4D) that might serve as a defence in response to a threat. This type of call is often accompanied by other aggressive behaviors such as body inflation or attacking toward an aggressor (Britton, 2001). These calls are noisy broadband signals (>15 kHz), with no clear harmonic structure.

### 1.3.2 Adults acoustic communication

Although it is established that all crocodylian species vocalize, the respective functions of all vocalizations have not been demonstrated experimentally. Most of the putative functions of adult vocalizations have been suggested by behavioral observations.

The most common and well-known long distance calls produced by adult crocodylians are the **bellows**, corresponding to loud, low-pitched and pulsatile booming roar (Garrick & Lang, 1977; Vliet, 1989). Most of the frequency energy is concentrated between 20 and 250 Hz, and the calls could spread over 160m in the air (Wang et al., 2007). Bellows are emitted by both sexes, mostly during courtship, and are part of a whole advertisement display (Campbell, 1973; Garrick & Lang, 1977). This complex behavior has been well documented particularly in alligators (Garrick & Lang, 1977; Wang et al., 2007). Males typically start the display by raising their heads and pulling their tails out of the water, a stereotypical position known as HOTA (“head oblique, tail arched”; Grigg & Kirshner, 2015). At the same time, they inhale deeply, loudly and audibly. Then they sink back into the water until only the head and tail remain on the surface and begin to produce the famous “water dance”. The water dance is actually the result of the production of infrasonic waves that project water droplets around the animal’s body. If the courtship of females starts in the same way by displaying the HOTA position, the water dance seems to be produced exclusively by males. In any case, there are no reports to date that females produce these infrasounds. The bellowing only follows the water dance. While its role is not clearly defined, it clearly stimulates and attracts other individuals, often causing a chorus of bellowing (Wang et al., 2007).

Crocodylians can also produce **hissing**, which can be considered as a defence or threat signal. It is often used during territorial interactions or nest defence, and can be followed by attempts to bite or retreat of the targeted individuals (Garrick & Lang, 1977; Britton, 2001).

A less documented type of adult call is the **grunt**, or the **maternal growl**, which are low-frequency sounds with a complex harmonic structure. During the first weeks after the hatching, the mother stays with her offspring. During this period, female American alligators have been seen to produce these low frequency rumble when guarding a pod (Hunt & Watanabe, 1982), resulting in the young being attracted to the female. Garrick & Lang (1977) recorded a female alligator grunt emitted in response to a juvenile contact call, as did Vergne et al. (2009) from a female spectacled caiman *Caiman crocodilus*. Mothers of black caiman *Melanosuchus niger* and Jacare caiman *Caiman yacare* have also been observed emitting this type of call with an



approach of their young as a response (Vergne, T. Aubin, P Taylor & N. Mathevon, unpublished observations; L. Papet, N. Grimault, N. Mathevon, unpublished observations). These calls might be used by females to maintain young together as a group and it is possible that this production is limited to the repertoire of Alligatoridae.

## **2. Perception and response to sounds by crocodilians**

### **2.1 Sound signals detection**

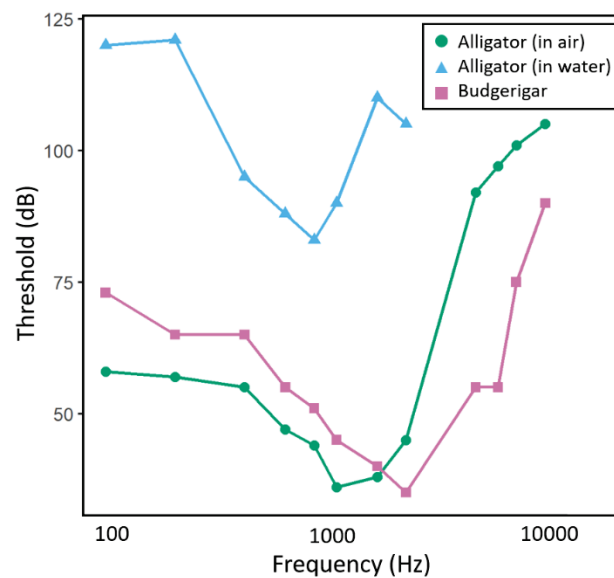
So far, it has been shown that globally the acoustic perception of sounds by crocodilians is very similar to that of birds (Vergne et al., 2009). Hearing sensitivity is therefore not so different from that of humans except that the range of hearing is restricted from a few hundred Hertz to just under 10 kHz (Dooling et al., 2000). Behavioral and physiological studies have provided hearing curves for different crocodilian species.

In general, crocodilians are sensitive for frequencies from 100 Hz to 8 kHz, with greater sensitivity in the 0.1-3 kHz (Figure 1.5; Wever & Vernon, 1957; Manley, 1970; Higgs et al., 2002; Vergne et al., 2009; Bierman et al., 2014). Their peak sensitivity is around 0.8-1.5 kHz, which corresponds to the dominant frequency bands of hatchling calls (Britton, 2001; Vergne et al., 2009). Early measurements of cochlear potential (Wever & Vernon, 1957) and measures of auditory brainstem responses yielded similar results in American crocodiles *Crocodylus acutus*, American alligators *Alligator mississippiensis*, and spectacled caimans *Caiman crocodilus* (Higgs et al., 2002). Similar auditory sensitivity for spectacled caimans has also been established by recording auditory fibre responses (Manley, 1970; Klinke & Pause, 1980).

When animals submerge, it is suggested that sound is perceived by bone conduction rather than through the tympanum. An air bubble is usually trapped under the ear lid and above the tympanum. But its artificial removal had no effect on auditory brainstem responses, suggesting an alternative pathway for sound perception under water.

Several recent studies support that crocodilians have efficient adaptations for directional hearing, which they use daily to analyse their sound environment (Bierman & Carr, 2015; Kettler & Carr, 2019; Papet et al., 2019). Their middle ear is coupled by air-filled cavities,

consisting of para-tympanic sinuses, which extend under and over the skull (Saunders et al., 2000; Witmer et al., 2008; Carr et al., 2016). Connections through this “interaural pathway” form a differential pressure receptor (Bierman et al., 2014) and contribute to enhance directional hearing, particularly for lower-frequency sounds (Dooling et al., 2000).



**Figure I.5.** Audiograms based on auditory brainstem response (data from Higgs et al., 2002).

They are able to locate a sound source with fairly high accuracy (Bierman & Carr, 2015; Carr & Christensen-Dalsgaard, 2015; Papet et al., 2020). Bierman & Carr (2015) trained juvenile alligators to localize a sound source. More recently, Papet et al. (2020) showed that Nile crocodiles *Crocodylus niloticus* are able to take advantage of both interaural time difference (ITD; corresponding to the difference in time of arrival of a sound at the two ears) and the interaural level difference (ILD) to accurately localize a sound source. Because higher frequency sounds are more likely to be filtered by the crocodile’s head, ILDs cues are more important and contribute greatly to the localization of sounds above 1.5 kHz (Bierman et al., 2014; Papet et al., 2019). Whereas ITDs are more accurate and might be used for low frequency sounds as they require precise neuronal temporal coding (Carr et al., 2009; Papet et al., 2019, 2020).

The well-developed directional hearing of crocodylians is a real advantage for the analysis of their sound environment. They can potentially segregate spatially distinct sound sources, and thus localize biologically relevant sounds (e.g. sounds from conspecifics or prey). However,



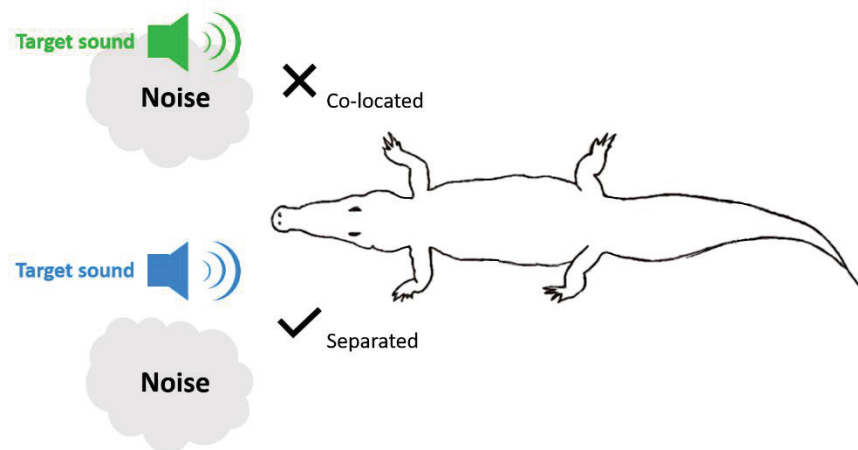


crocodilians are facing a mass of auditory information in their complex sound environment (Figure I.6). Relevant sound sources can be mixed and masked by irrelevant sounds. This constant noise can drastically reduce their sound detection and thus sound localization performance, a critical aspect especially for very young crocodilians for whom acoustic communication is essential to their survival.



**Figure I.6.** Illustration of a complex sound environment. In this amphibious environment, crocodilians may be exposed to a variety of sound sources, both biotic (e.g., the calls of conspecifics or chorusing frogs) and abiotic (e.g., the sound of waterfalls, wind). Multiple sounds can form a background noise that masks signals of biological interest and alters the acoustic communication of crocodilians. Once a sound is detected among the noise, the crocodilians must then identify it, or even classify it into a functional category, and then decide if this signal is worthy of interest or not.

Derived from their localization abilities, Spatial Release from Masking (SRM) is a mechanism that can be used by crocodylians to overcome this problem of sound localization in noise. The basic principle is that sound detection is enhanced when a relevant sound (called the target) is spatially separated from a noise source (called the masker) compared to when the two sound sources are co-located (Figure I.7; Litovsky, 2012; Saberi et al., 1991). This ability is used by a wide range of animals, including birds (see *part III.1*). It is therefore likely that crocodylians may also benefit from spatial unmasking to detect target sounds masked by the environmental noise.



**Figure I.7.** Basic principle of Spatial Release from Masking. Sound signal detection threshold is lowered when the signal is separated from the noise source.

## 2.2 Conspecifics information identification and decision-making

Juvenile crocodile calls possess a common basic acoustic structure, which varies along a continuum to grade into one or another call type (Vergne et al., 2012). Furthermore, a growing body of research suggests that all crocodylian species may share this common vocal repertoire, which may not be species-specific. This suggests a fine-grained coding of information that is universal across species. This idea was established with Campbell (1973) who showed that American alligators *Alligator mississippiensis* were responsive to the calls of *Caiman crocodylus*, *Melanosuchus niger* and also American crocodiles *Crocodylus acutus*.

Juvenile Nile crocodiles *Crocodylus niloticus* also responded equally to contact calls of their own species compared to calls of Black caimans *Melanosuchus niger* and Spectacled caimans



*Caiman crocodilus* although the acoustic parameters of the calls showed significant inter-specific differences (Vergne et al., 2012). In this study, the authors also showed that the slope of frequency modulation is the key parameter responsible for identifying a juvenile contact call, suggesting that information is universally encoded through this parameter.

Concerning distress calls, their duration and the acoustic parameters related to their fundamental frequency seem to vary with the age just after hatching (Vergne et al., 2007). The youngest individuals - thus the smallest - produce the highest pitched calls, but the slope of the frequency modulation remains stable. This result was confirmed by Chabert et al. (2015) who showed that the fundamental frequency of distress calls decreased with increasing body size in Nile crocodiles *Crocodylus niloticus*, American alligators *Alligator mississippiensis*, Spectacled caimans *Caiman crocodilus*, Morelet's crocodiles *Crocodylus moreletii* and Orinoco crocodiles *Crocodylus intermedius*. This information appeared to be of interest of caring parent who were more responsive to the distress calls of smaller (and generally younger) juveniles. Crocodilian parents might therefore perceive the distress encoded in the calls of juveniles, and adapt their response to help the youngest who are most exposed to predation.

Apart from these studies, where the experiments have remained fairly basic and poorly replicated, it is still unclear how information is encoded and how auditory information is decoded and perceived by young crocodilians. Once animals are able to detect and localize sounds, the question is which sound is relevant and interesting relative to another, and what makes a sound interesting enough for the crocodile to decide to respond. And for example, basically, how do they identify a call made by a conspecific and how do they identify it as a specific type of call (Figure I.6)?

From a cognitive psychology perspective, categorization is a cognitive process that can aid in the analysis of a sound scene by determining which stimuli or events can be grouped together in a coherent manner. Acoustic stimuli can be grouped into the same category on the basis of shared physical characteristics as well as on the basis of shared function (Zentall et al., 2002). While juvenile contact calls and distress calls seem to be graded into one another, with no clear acoustic boundary, they are perceived distinctly by youth and adults, giving them a well distinct function (at least the female) (Britton, 2001; Vergne et al., 2009). This implies that categorical perception may be used by crocodilians to classify their own call type, and yet also to classify the majority of sounds in their environment. If this is the case, an important question remains

to be able to identify acoustic features that could be used to create acoustically coherent objects and to extract them from the complex auditory scene (Hulse, 2002).

### **2.3 Taking advantage of the sound environment: perceive the others**

When analyzing their sound environment, crocodilians can also take advantage of information conveyed by sounds produced by other animals present, such as prey, predator or impending danger (Figure I.6). No studies have examined crocodilian's perception of the sound environment in a broad sense, including their perception of other animals. Studies of sound perception in crocodilians have focused on signals specific to the crocodilian vocal repertoire, or on the use of synthetic signals for experiments that include conditioning (Bierman & Carr, 2015; Papet et al., 2020).

This type of study on interspecific communication has become more importance in recent years. They have focused on several species but remains restricted to mammals and birds. This is a subject that interested me during my PhD, and on which we do not have a specific bibliographic base for crocodilians. To know if crocodilians can obtain information about other animals through their vocalizations. Are they able to perceive and use the acoustic information that other species use among themselves and that are not intended for them. To focus on a specific aspect of interspecific communication, we already knew that crocodilians discriminate behaviorally between distress calls and contact calls: they are only receptive to the distress calls of juveniles. Are they therefore also able to perceive the distress encoded in the vocalizations of other species?





## II. General methods

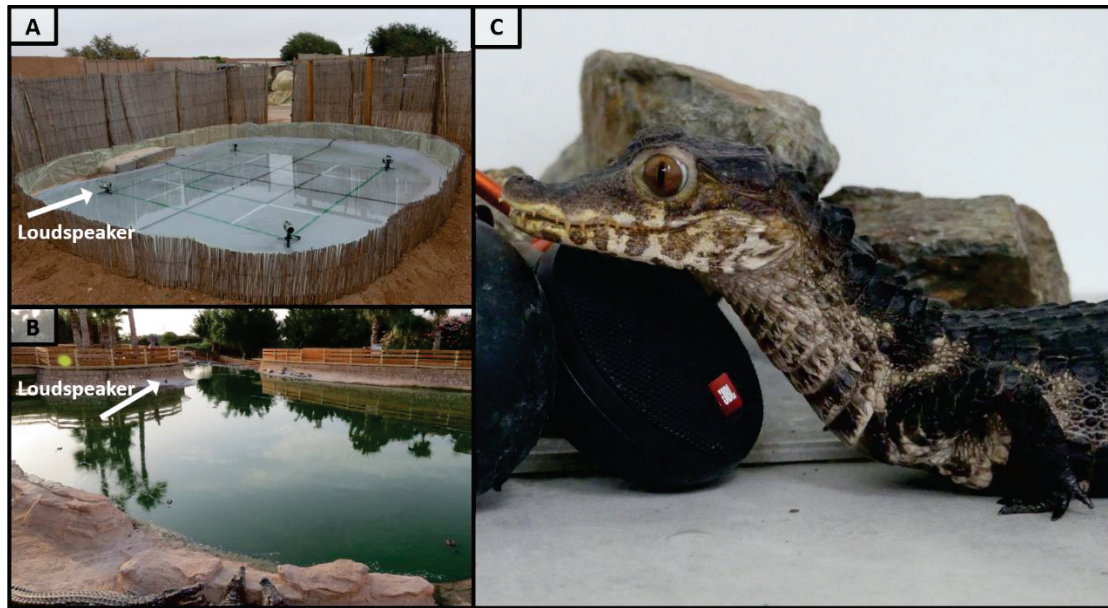
### 1. Playback experiments

Two experiments were conducted in Crocoparc Agadir in Morocco. The Crocoparc Agadir is a huge zoologic park housing no less than 350 Nile crocodiles (*Crocodylus niloticus*), as well as several specimens of Broad-snouted caimans (*Caiman latirostris*). Naturally present until the late 70's in the gueltas of the Oued Draâ in Morocco, the crocodiles have gradually disappeared due to dryness but also to excessive human hunting. The park was created in 2015 for the purpose of conservation of the species and population education. I conducted different experiments on juvenile and adult Nile crocodiles. In all cases, the animals were completely free to move throughout the experiments. The experiments basically consisted of playing acoustic stimuli to the animals (single individuals for the experiments involving juveniles, or groups of individuals for adults) through loudspeakers and recording/scoring their behavioral response.

A part of the first study (*part III.2.2*) as well as of the second one (*part IV.1*) were performed on juvenile Nile crocodiles placed individually in turn in a dedicated experimental pond (Figure II.1A). The animals were tested at nightfall and only once. As the juveniles were in turn isolated from the rest of the group and placed in a novel environment, we placed them at least 2 hours before the start of the experiment in the experimental pond without any stimuli. This acclimatization period was intended to give them time to explore their new environment and to considerably reduce their stress. Once the experiment was completed, the young crocodile was placed back with the rest of the group and identified so that it would not be tested again.

The last playback experiment (*part V*) was performed on adult Nile crocodiles moving freely in their holding tanks (Figure II.1B). The difficulty here was to get the crocodiles used to the speakers placed in their enclosure. To do this, we placed the speakers in the intended locations for the experiment for several hours, a few days before the experiments. The second challenge was to test a group of individuals rather than a single one, due to the high density of crocodiles per pond. I therefore chose to quantify the behavioral response of crocodiles in this experiment as the percentage of individuals in the group that responded. However, we cannot rule out (and this is probably the case) that some individuals responded by imitation rather than by their own initiative.





**Figure II.1. A, B:** Experimental setup used in the playback experiments performed in Crocoparc Agadir (presented in *part III.2.2, IV.1* and *V*). **C.** Example of a conditioning session carried out in the laboratory (see *part II.2.3*): the caiman was trained to come spontaneously to the loudspeaker when it emitted a sound signal.

## 2. Laboratory experiments

The laboratory experiments concern one of the experiments presented in *part III.2.3* as well as the experiments presented in *part IV*. I will speak here only about the methodology set up in order to carry out the experiments presented in *part IV*, namely the experiments carried out with the Cuvier's dwarf caimans since it is me who had the responsibility of these animals throughout my thesis.

### 2.1 Generalities

I worked with 6 Cuvier's dwarf caimans *Paleosuchus palpebrosus* (also called Cuvier's smooth-fronted caimans), all from the same clutch. They were born in the Parc zoologique de Paris (Vincennes, France) in 2017 and arrived at the ENES lab in 2018 (under the laboratory ethic agreement n° D 42-218-0901 and the certificate of capacity of Nicolas Boyer). They were never included in any experiments or conditioning protocols until I arrived in 2019.

Dwarf caimans are the smallest of the living crocodylians. Widespread across northern South America, the dwarf caimans are found across a large range of tropical habitat such as flooded forest including rivers, lakes, streams or swamps (Stevenson, 2019). As in other crocodylians, dwarf caimans exhibit parental care and has been observed caring for their young for 21 months after hatching (Campos et al., 2012) which makes them interesting for the study of acoustic communication, especially in juveniles.

The challenge in laboratory is to reproduce the animal's natural environment. Even when working with dwarf species, crocodylians grow to be very large (adult male reaches an average size of 2m) and therefore have substantial space and logistic requirements. Their semi-aquatic lifestyle and their preferential habitat required the installation of a large enriched enclosure with access to tanks filled with water. The ambient temperature was stabilized around 29°C. The water temperature was constant around 26°C to allow the animals to reduce their body temperature and hot spots at 32°C were set up with UV-lights (also essential for the synthesis of vitamin D<sub>3</sub> which allows the metabolism of calcium and phosphorus). The rooms were kept under a circadian light cycle (12:12). Dwarf caimans are dietary generalists which makes them easy to feed. When they first arrived as very young, the animals were fed with small fish, frogs and insects, before moving on to small mammals such as mouse foetuses. To ensure the well-being of the animals and their good cohabitation, the enclosure was filmed 24 hours a day and routine examinations were carried out every month, including weighing, measuring, checking teeth and the general condition of the animals. The routine exams imply that the caimans have been accustomed to being handled regularly, as we will see below.

## **2.2 General principle of conditioning**

My laboratory experiments were all based on classical conditioning of the animals (Figure II.1C). The classical conditioning is a paradigm set up by Ivan Pavlov at the beginning of the 20<sup>th</sup> century. He showed that if the presentation of food to a dog was always accompanied by the sound of a bell, the dog would react to the bell as if it were food. By studying the dog's salivary response, he established that salivation in front of food was an unconditional response, whereas salivation in response to a bell sound was a conditional response since it depended on the prior coupling of the food and the bell sound (Pavlov, 1927). Therefore, food was considered an unconditional stimulus while the sound of the bell was considered a conditional stimulus.





Classical conditioning in general is therefore the result of an event organised by the experimenter, between an unconditional stimulus (food) and an external conditional stimulus not yet related to food (which for us will be a sound stimulus). After a certain number of couplings and thus after a certain learning period, the conditional stimulus alone is sufficient to trigger a foraging response in the animal. During this conditioning phase when the animal learns to couple the two stimuli, the presentation of the unconditional stimulus (food) following the conditional stimulus (sound stimulus) is said to reinforce the animal's conditional response. The food thus serves as a positive reinforcer in the conditioning process. Negative reinforcers can also be used in the conditioning process, but these are more difficult to implement because of their detrimental effect on long-term motivation and animal welfare more generally.

The reinforcer should be chosen for its motivational quality. For example, food is a good reinforcer only if the animal is hungry. The speed of acquisition of conditioning (or coupling between unconditional and conditional stimuli) varies greatly between species, ages, and the stimuli used and circumstances. As much as to say that it is not easy and widespread with crocodilians at all.

## **2.3 Conditioning applied to crocodilians**

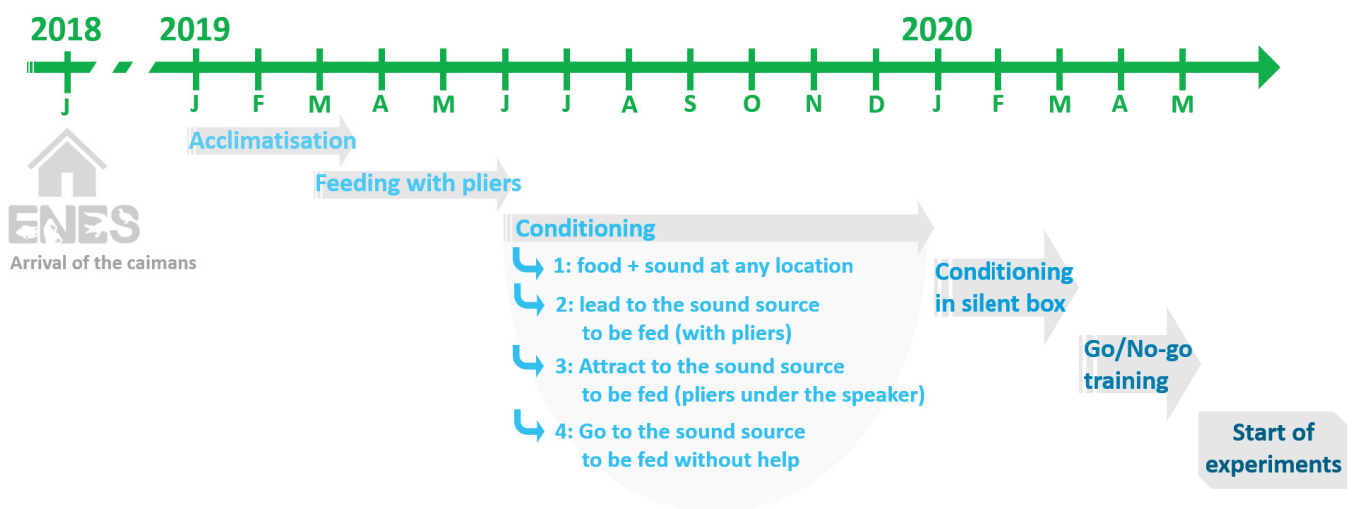
### **2.3.1 Acclimatisation**

The conditioning of the young Dwarf caimans was long and tedious (Figure II.2. Maybe because of their small size, and therefore their greater susceptibility to predation, dwarf caimans are particularly fierce and nervous, which made them very difficult to acclimatise and handle.

When I arrive, the animals were frightened by the human presence and displayed a set of agonistic behavior (as defined in Brien et al., 2013) towards me and the animal keeper (from the most frequent to least frequent): body inflation accompanied or not by threat calls, mouth agape posture (immobile with mouth opened wide), head raised high posture (immobile with upward extension of the front two limbs pushing the head and chest high out of the water), side head-strike (head is thrust sideways in to an opponent while the mouth is either open or closed), tail-wagging (undulation of the tail from side to side in either a gentle sweeping motion or rapid twitching, repeated several times) and even bite attempt. Setting up their handling was also difficult (and even more so over time as the animals grew) due to their extensive and prominent

osteoderms contained in their skin, providing them a heavy protective layer but being very uncomfortable when manipulating.

Because of their fear and aggressiveness, the conditioning was preceded by a long acclimatisation phase for the animals, consisting on visiting them several hours each day. Once the animals were less fearful of my presence (i.e. once they had reduced their aggressive or offensive behavior in my presence), I started to get them used to being fed with pliers at a frequency of 2 feedings per week.



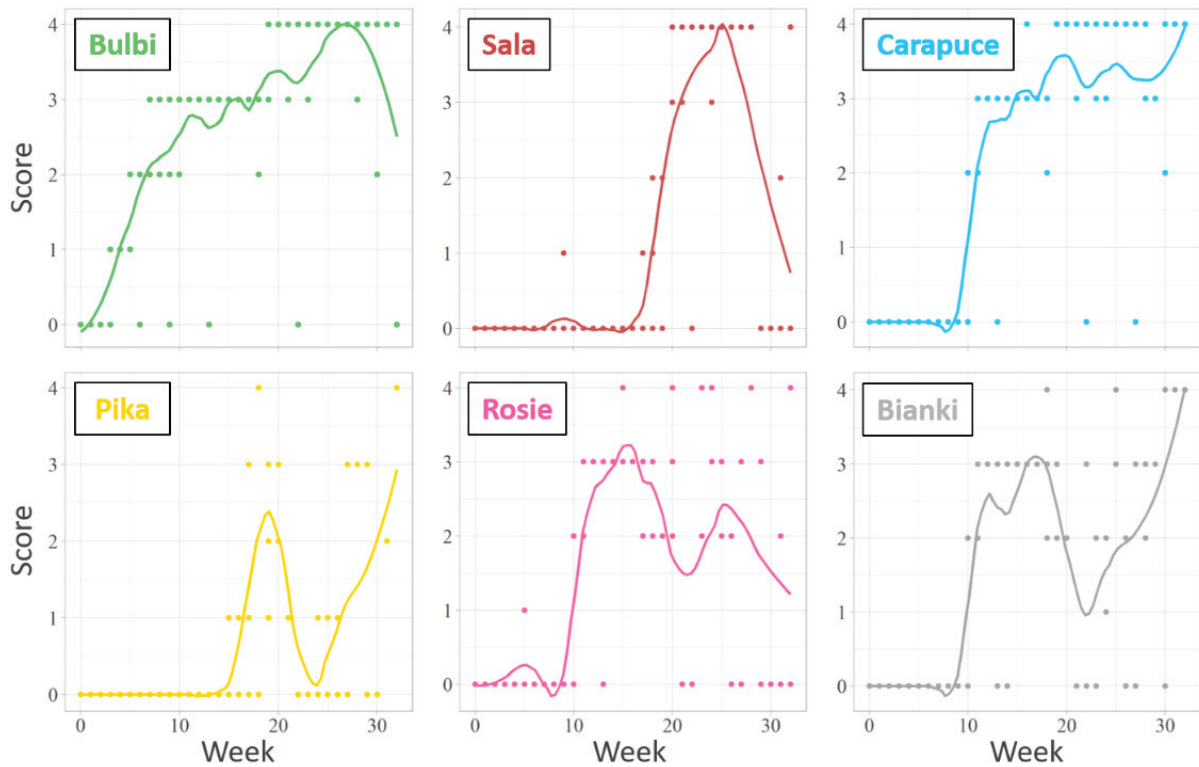
**Figure II.2.** Timeline of conditioning progress of the caimans in the laboratory.

### 2.3.2 Conditioning

The first step in the conditioning process was then to put a loudspeaker playing continuously a series of neutral sound signals (buzz series) during these feeding sessions. This stage was probably the longest stage in the conditioning process as it took almost 6 months for the caimans to accept to take the food with the pliers and eat it in my presence.

The conditioning consisting on the association of the sound source with the food then began in earnest by requiring the caimans to move towards the sound source to get their food. At first, they were guided by the pliers with food at the end until they reach the loudspeaker, then gradually I just placed the pliers with food just under the loudspeaker to attract them. After several more months, the caimans have associated the sound source with food and are able to come to the active speaker without the help of the pliers (Figures II.2 and II.3).





**Figure II.3.** Conditioning curves of the 6 Dwarf caimans. The behavior of the animals during the conditioning sessions was scored as indicated in Figure 9.

At this point I then changed the sound signal (the buzz series) to the signal I would use in my experiment, which was a crocodile contact call. The animals continued to come towards the sound source after this change, probably because they transferred what they learned to any sound signal coming out of the loudspeaker, but also perhaps, and this we will never know, because contact calls naturally generate (at least with wild animals) an approach behavior of juveniles.

Once the caimans had assimilated this task, the weekly feeding sessions only took place in an experimental room consisting of a silent box in which a pool filled with water was placed. The animal was then only individually conditioned. A loudspeaker was placed at each corner of the pool, and was associated with a custom-made food delivery system. I could control all the elements (sound emission and food delivery) so as not to interfere with the animal's response. This step, which involved catching and moving the animal, added stress that set back the conditioning process for some of the individuals for a few weeks.

### 2.3.3 Go/No-go conditioning

In order to perform the categorisation experiments presented in *part IV.1*, the animals then had to be trained in a Go/No-go paradigm. In addition to having to respond to the crocodile call, they then had to learn not to respond to a frog call, and thus not to generalise their approach behavior to any sound stimuli. Interestingly, some individuals did not respond to the frog call from the first go/no-go session while others had to learn to disengage from the signal by associating the frog call with the absence of food reward. Overall, the go/no-go learning process took an average of 10 weeks.

## 2.4 General feedback

In general, conditioning took a very long time (> 8 months) to set up because of the initial stress of the animals to my presence or the presence of new objects. When they are not used to human presence at a very young age, the dwarf caimans seem to develop a very shy, fearful behavior and feel threatened when human approach. This behavior differs from that we could have observed in the laboratory in other less cryptic and larger species such as the Nile Crocodile.

Another important aspect to take into account during this conditioning process was the fact that the animals have a bradymetabolism, thus a slow digestion, and cannot be motivated by food rewards over several days. The conditioning steps as well as the experiments had to take place only at a rate of one day per week in order to respect their feeding rhythm, thus spreading over several months. During the passage in the experimental room, this implied a weekly capture of the animals, constraining and stressful for both the animal and the experimenter. Although the animals became accustomed to the regular capture, rest periods were observed every month and a half to ensure their well-being.

Subsequently, the animals assimilated the sound source with the food more or less quickly depending on the individual (Figure II.3). One individual was much quicker to condition than the others (*Bulbi*, Figure II.3). Group structure seemed to have a significant influence on learning. Some individuals in the group seemed to imitate the behavior of the earliest individual, even associating themselves the sound signal with food (social learning by copying and/or accentuation). In this case, the group effect was beneficial for learning. This was also seen in the fact that one individual, isolated for health reasons, made a jump in conditioning progress once he joined the rest of the group, suggesting at least some emulation due to the presence of



other individuals. Finally, one of the six caimans never managed to pass the go/no-go conditioning and was excluded from the experiments.

In addition, a hierarchy gradually developed in the group. The most dominant individuals inhibited the approaches to the speaker of the less dominant individuals, by showing aggressive behavior. This was not really surprising since dwarf caimans are considered as mildly tolerant towards conspecifics (Brien et al., 2013). In this case, the group had a negative effect on learning, to the point of requiring the group to split into two separate groups.

Generally, the most dominant and aggressive individuals were also the boldest and exploratory in the experimental room, but also the most problematic with the experimental set-up. The set-up had to be revised several times until it was optimal so that the animals could not climb on the speakers or reach the food hidden in the reward system.

These observations as well as my daily observations support the idea that the dwarf caimans, at least the 6 that we had here, have their own personality, which is reflected in their behavior towards other individuals, towards new tasks involving new objects and new places, or towards their more or less aggressive or fearful behavior towards the experimenter (Table II.1).

Outside the periods of conditioning or experimentation, I set up a feeding in the form of training, consisting in making the animal move from a point A to a point B with the help of a stick to a feeding area. This training allowed me to better know my animals individually but also to overcome the problem of dominance and aggressiveness during common feedings, by targeting only one individual to move and feed at a time.

	Sex	Size	Conditioning speed	Training performance	Catching	Tolerance to conspecifics	Area exploration
<b>Bulbi</b>	F	Medium	Fast	Excellent	Easy	Yes	High
<b>Sala</b>	F	Medium	Medium	Low	Easy	Yes	Low
<b>Pika</b>	M	Small	Low	Medium	Difficult	No	Low
<b>Rosie</b>	M	Small	Good	Excellent	Difficult	No	Medium
<b>Carapuce</b>	M	Large	Good	Good	Medium	Medium	High
<b>Bianki</b>	M	Large	Very low	Low	Medium	Medium	Low

**Table II.1.** General observations on the 6 dwarf caimans included in the conditioning procedure.

Whether it is for training outside of the experimental periods or for conditioning, the redundancy of the tasks, with a multisensory aspect (visual, sound, olfactory, always involving the same clamp, or the same stick, the same experimenter, etc.) and at a fixed day and time, have been the key to work with these caimans.

### 3. Timeline of the experiments

The articles presented below follow the framework presented in the second part of the introduction and their purpose is to answer the following questions:

**Part III:** Do crocodilians are able to use spatial release from masking to improve sound detection and localization in noisy environment? We answer this question through a multi-contextual approach combining playback and laboratory crocodile conditioning experiments. This study was published in *Communications Biology* (Thévenet et al., 2022b).

**Part IV:** Once crocodiles are able to detect and localize sounds, what makes a sound interesting enough for the crocodile to decide to respond to it? Do crocodilians can categorize an auditory continuum? Again, to answer these questions, I combine zoo playback experience and extensive laboratory studies with caimans conditioning. I also present a complementary experiment aimed at defining the discrimination performance of signals belonging to the continuum. For this I used the principle of pupillometry, in which pupil dilation can be an indicator of signal discrimination in an oddball paradigm.

**Part V:** Female crocodilians are very sensitive to the distress calls of juvenile crocodiles. As top predators with good auditory abilities, are crocodilians also able to perceive distress encoded in the vocalizations of other species and take advantage of this information? To explore this question, we conducted playback experiments on large groups of Nile crocodiles in zoos.





# III. Sound detection in noise

## Spatial Release from Masking in crocodilians

Article published in the *Communications Biology* journal (<https://doi.org/10.1038/s42003-022-03799-7>; Thévenet et al., 2022b)

**Thévenet, J.<sup>1,2\*</sup>, Papet, L.<sup>1,2\*</sup>, Campos, Z.<sup>3</sup>, Greenfield, M.<sup>1,4</sup>, Boyer, N.<sup>1</sup>, Grimault, N.<sup>2‡</sup>,  
& Mathevon, N.<sup>1‡</sup>**

\* Co-first authors ‡ Co-last authors

<sup>1</sup> Equipe de Neuro-Ethologie Sensorielle ENES / CRNL, CNRS, Inserm, University of Saint-Etienne, Saint-Etienne, France

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

<sup>3</sup> Wildlife Laboratory, Brazilian Agricultural Research Corporation EMBRAPA, Corumbá, Brazil

<sup>4</sup> Department of Ecology and Evolutionary Biology, University of Kansas, Lawrence, KS 66045, USA

## Abstract

Ambient noise is a major constraint on acoustic communication in both animals and humans. One mechanism to overcome this problem is *Spatial Release from Masking* (SRM), the ability to distinguish a target sound signal from masking noise when both sources are spatially separated. SRM is well described in humans but has been poorly explored in animals. Although laboratory tests with trained individuals have suggested that SRM may be a widespread ability in vertebrates, it may play a limited role in natural environments. Here we combine field experiments with investigations in captivity to test whether crocodilians experience SRM. We show that 2 species of crocodilians are able to use SRM in their natural habitat and that it quickly becomes effective for small angles between the target signal source and the noise source, becoming maximal when the angle exceeds 15°. Crocodiles can therefore take advantage of SRM to improve sound scene analysis and the detection of biologically relevant signals.



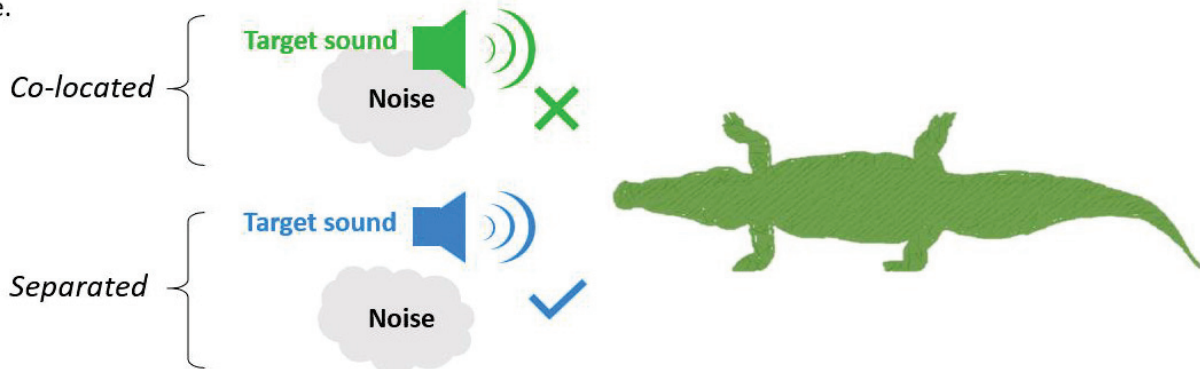


# Graphical abstract

## Spatial release from masking in crocodilians

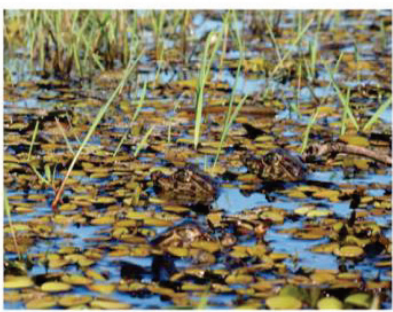
Do crocodilians are able to use spatial release from masking to improve sound detection and localization in noisy environment?

Constant noise can significantly reduce sound detection, a critical aspect for young crocodilians for whom acoustic communication is essential to their survival. **Spatial Release from Masking (SRM)**, the improvement of sound detection when a relevant sound (the target) is spatially separated from a noise source (the masker) compared to situation where both sound source are co-located, could be used by crocodilians to overcome this issue.



### 3 experiments, 3 contexts

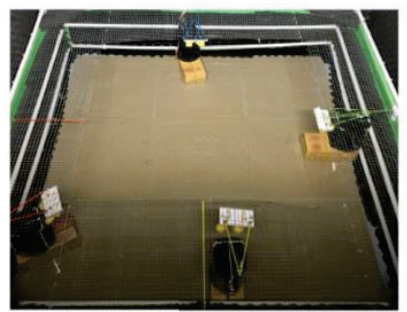
Naive Yacare caiman in wild  
Mother response to distress calls



Naive Nile crocodile in Zoo  
Juvenile response to contact calls



Conditioned Nile crocodile in lab  
Response to synthetic buzz



Amount of Spatial Release from Masking

(= difference in auditory threshold between the co-located and the separated configuration)

SRM =  
3.6 dB

SRM =  
6.8 dB

SRM =  
2.8 dB

Optimal for angles > 15°



### Conclusion

Crocodilians detect sound signals better when the target source is spatially separated from the masking noise source, suggesting that these animals use SRM in their daily lives to improve sound scene analysis and the detection of biologically relevant signals. The differences in amount found can be partially explained by differences in ethological contexts involving different motivational state, stress, and trade-offs.

# 1. Introduction

Animals that use acoustic signals to communicate often develop strategies for optimizing information transfer in noisy soundscapes (Lohr et al., 2003; Brumm & Slabbekoorn, 2005; Mathevon & Aubin, 2020; Duquette et al., 2021; N. Lee et al., 2021; Gomes et al., 2021). Emitters may increase Signal-to-Noise Ratio (SNR) by raising signal intensity (Lombard effect; Manabe et al., 1998), by shifting signal frequency to avoid overlap with the noise frequency bandwidth (e.g. in great tits *Parus major* and zebra finches *Taeniopygia guttata*; (Slabbekoorn & Peet, 2003; Villain et al., 2016), by using signal redundancy (Lengagne et al., 1999; Aubin & Jouventin, 2002; Foote et al., 2004), or by choosing emission posts and behavioral postures that optimize signal transmission (e.g. songposts; Dabelsteen et al., 1998; Mathevon et al., 2005, 2008). At the other end of the communication chain, receivers may choose strategic posts and behaviors that improve signal reception and facilitate auditory computation in noisy environments (e.g. hearing posts in songbirds (Mathevon & Aubin, 1997; Nemeth et al., 2006). When listening in noise, spatial cues such as Interaural Time Differences (ITD) and Interaural Level Differences (ILD) play an important role in improving signal detection, source localization, and information decoding (Knudsen & Konishi, 1979; Blauert, 1997; Schnupp & Carr, 2009; Carr et al., 2016). Spatial Release from Masking (SRM) refers to the process where the auditory system of listeners uses these directionally dependent cues to segregate the signal of interest (target) from competing sounds (maskers; Bee & Micheyl, 2008; Litovsky, 2012). According to SRM, signal reception is better when the signal source is spatially separated from the noise source than when both signal and noise sources are co-located in the environment (Saber et al., 1991; Litovsky, 2012).

SRM has primarily been investigated in humans. The seminal study by Saber et al. demonstrated that SRM is efficient in both the horizontal and vertical planes in our species (Bronkhorst, 2000). SRM has also been found in a few other mammal species: ferrets *Mustela putorius* (Hine et al., 1994), cats *Felis catus* (Wakeford & Robinson, 1974), big brown bats *Eptesicus fuscus* (Sümer et al., 2009), harbor seal *Phoca vitulina* and sea lion *Zalophus californianus* (Holt & Schusterman, 2007). In birds, SRM enhances the detection of pure tones masked by a broadband noise in budgerigars *Melopsittacus undulates* (Dent, 1997), and the detection of bird songs in a song chorus in both zebra finches *Taeniopygia guttata* and budgerigars (Dent et al., 2009). SRM has been investigated in amphibians (northern leopard frogs *Rana pipiens pipiens* (Ratnam & Feng, 1998), Cope's gray treefrog *Hyla chrysoscelis*



(Bee, 2008; Bee & Vélez, 2018; Caldwell et al., 2016; Rocchi et al., 2017; Greene et al., 2018), showing better detection and discrimination of conspecific calls masked by noise when the two sources are spatially separated (Bee & Micheyl, 2008; Nityananda & Bee, 2012). Finally, SRM has also been found in two crickets (*Paroecanthus podagrosus* and *Diatrypa sp.*), where it improves the detection of natural conspecific song against the ambient noise of the rainforest (Schmidt & Romer, 2011). Notably, the fly *Ormia ochracea* is the only known animal species which seems not able to benefit from SRM (N. Lee & Mason, 2017). In addition to sound communication in air, SRM has also been found in underwater communication with bottlenose dolphins *Tursiops truncatus* (Popov et al., 2020).

Although SRM could appear as a widespread ability to increase the detection of sound signals against masking noise, it has yet been investigated in a limited diversity of experimental approaches and situations. All previous studies investigating SRM in animals have been performed in very controlled conditions in the laboratory or captive environments (Wakeford & Robinson, 1974; Hine et al., 1994; Dent, 1997; Bee, 2007; Holt & Schusterman, 2007; Sümer et al., 2009; Dent et al., 2009; Nityananda & Bee, 2012; Caldwell et al., 2016; Popov et al., 2020). There has been no field investigation with animals freely behaving in their natural habitat. This is a serious limitation: it cannot be ruled out that SRM is a laboratory artefact with a limited role in the field. Indeed, in the field, animals are exposed to a wider and more realistic range of situations, e.g. in terms of head position relative to the sound source and noise. Testing their SRM abilities in field condition would certainly provide a more realistic picture. Moreover, all studies performed in vertebrates (except one with treefrog; Nityananda & Bee, 2012) have been based on conditioning experiments where animals were trained to locate sound sources (Go/No-Go experiments; Wakeford & Robinson, 1974; Hine et al., 1994; Dent, 1997; Bee, 2007; Holt & Schusterman, 2007; Dent et al., 2009; Sümer et al., 2009; Caldwell et al., 2016). While Go/No-Go experiments may limit the variability of the tested subjects' motivation, an intensive training combined with laboratory conditions is likely to change the ability of subjects to perform SRM compared to natural field conditions. Strikingly there has been no study on SRM combining different experimental approaches, in both controlled and natural settings. In spite of its tremendous utility for sound scene analysis in the daily life of animals, SRM thus remains a poorly investigated phenomenon.

In the present study, we investigated SRM in crocodylians. These animals may indeed be ideal subjects for studying SRM in various conditions for the following reasons. First, they are relatively immobile, which allows us to conduct these acoustic experiments in the field with a

precision and a control of initial conditions usually restricted to laboratory experiments. Second, they do actively use acoustic communication during their social interactions (Grigg & Kirshner, 2015; Vergne et al., 2009), where the detection of signals could be critical. Mature embryos vocalize to synchronize hatching and promote maternal care (Vergne & Mathevon, 2008). Juveniles emit contact calls ensuring group cohesion, and distress calls inducing maternal protection (A. L. Vergne et al., 2012; Sicuro et al., 2013). Adult males of most species attract females and repel competitors by producing a repertoire of vocalizations (bellows, grunts) as well as low frequency sounds through the vibration of their whole body (Todd, 2007), while females emit grunts to attract their young (Vergne et al., 2009). Third, crocodylians spend most of their active life cruising at the interface of air and water. In this amphibious environment, they can be exposed to various sources of noise, either biotic (e.g. chorusing frogs) or abiotic (e.g. waterfall noise, anthropogenic noise such as boats). This noise may mask crocodylians' vocalizations and may thus impair their acoustic communication. The receiving individual must discriminate the signal of interest against non-relevant masking sounds, and SRM could represent a valuable ability. Moreover, the head morphology of crocodylians enables them to acquire reliable localization cues from sound sources propagating in the air even when only a small part of their head is above the air-water interface (Bierman et al., 2014; Papet et al., 2019). In a previous study, we found that crocodiles may use both Interaural Level Differences cues and Interaural Time Differences cues to accurately locate the spatial direction of a sound source (Bierman & Carr, 2015; Kettler & Carr, 2019; Papet et al., 2020). However, the radically different acoustic impedances of air and water prevent most of the acoustical energy from entering the water and thus removes part of the acoustical difference between right and left ears (Bierman et al., 2014; Papet et al., 2019).

Here we demonstrate that crocodylians cruising in water use SRM to detect target sounds against a noisy background both propagating in the air. We used three different experimental paradigms to explore this ability. First, we examined whether adult crocodylians (*Caiman yacare*) use SRM in natural conditions by performing field experiments in the Pantanal, Brazil. We challenged naive caiman mothers while they were caring for their young by mimicking a situation where an isolated nestling was emitting distress calls (Vergne et al., 2007). We then tested whether these SRM abilities are already present in young crocodylians with experiments in a zoo, where we assessed the response of naive young Nile crocodiles *Crocodylus niloticus* to the playback of contact calls (Vergne et al., 2012). Jacare caimans and Nile crocodiles are two representatives of two of the three extant groups of crocodylians that differentiated during the Cretaceous: the



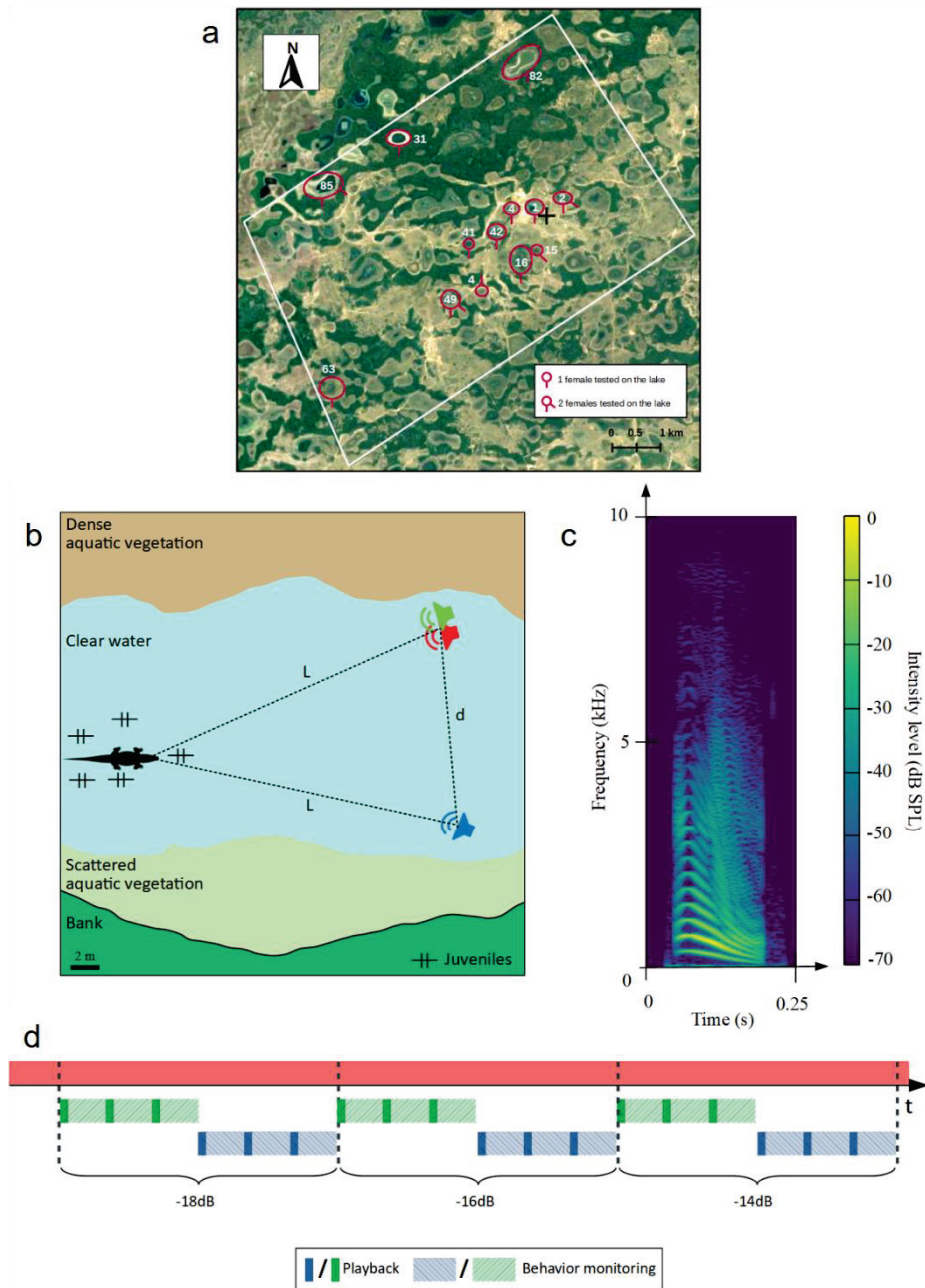
Alligatoidea and Crocodyloidea respectively (the third group being the Gavialoidea). Finally, we tested whether SRM functions with non-biological signals in the laboratory by training juvenile Nile crocodiles to identify a synthesized sound from a masking noise using Go/No-Go experiments. In these three experimental situations, we evaluated the ability of the tested individuals to detect the source of the target signals as a function of the location of the background noise source.

## 2. Results

### 2.1 Spatial Release from Masking by adult crocodilians

This first experiment was conducted on wild adult female yacare caimans *Caiman yacare* (Pantanal, Brazil, Figure III.1). For each female (N = 16), we played back a broadband noise ("noise source", emitted at 83 dB SPL, unweighted) simultaneously with a series of distress calls recorded from young juveniles ("target signal"; relative intensity to the noise in the range [-20, 0]dB; calls recorded from 3 week old individuals, unfamiliar to the tested females, see Figure III.1c for the spectrogram of one call). In this experiment, distress calls were selected to optimize a behavioral response from the female toward the loudspeaker (Vergne et al., 2009). As illustrated on Figure III.1b, the two loudspeakers emitting the "noise source" and the "target signal" were either side-by-side ("co-located condition") or spaced apart ("separated condition"; mean separation angle between the female, the noise source and the distress calls source = 18°, min-max = 4-44°, Raw data in appendix Supplementary Figure III.1a). Each female was successively challenged with several co-located and separated target signals with various Signal-to-Noise Ratios (SNR) and separation angles (16 females tested, mean number of trials per female = 6.5±5; Figure III.1d, see in appendix Supplementary Table III.1 for details on the signals played back to each female).

At the beginning of each playback, the loudspeakers were at approximately 20 meters from the tested female. We rated the female's response to playback according to a 0-4 level behavioral scale (score for no reaction = 0; head or body movement not in the direction of the target loudspeaker = 1; head movement towards the target loudspeaker = 2; displacement on a distance less than 1 body length towards the target loudspeaker = 3; displacement on a distance more than 1 body length towards the target loudspeaker = 4). We compared the behavioral reactions



**Figure III.1. Field experiments on female Jacare caimans (Experiment 1).** **a** Cartography of the field work area (white rectangle = border of the Nhumirim reserve; black cross = field station; 18°59'16.1"S 56°37'08.8"W). We conducted the experiments in the lakes surrounded by red circles. The number of red hyphens indicates the number of females tested on the same lake (1 or 2 individuals). **b** Schematic representation of the experimental design. The distance between the loudspeakers ([6.5, 19]m) was always lower than the distance between the female and the loudspeakers ([12, 50]m). **c** Spectrogram of a distress call from a young Jacare caiman. **d** Timeline of an experiment. The masking noise is emitted continuously (red solid line). The target signals (with different SNR) are emitted either by the co-located loudspeaker (green solid line) or by one of the separated loudspeakers (blue solid line). The light dashed lines represent the behavior recording following the target emission.

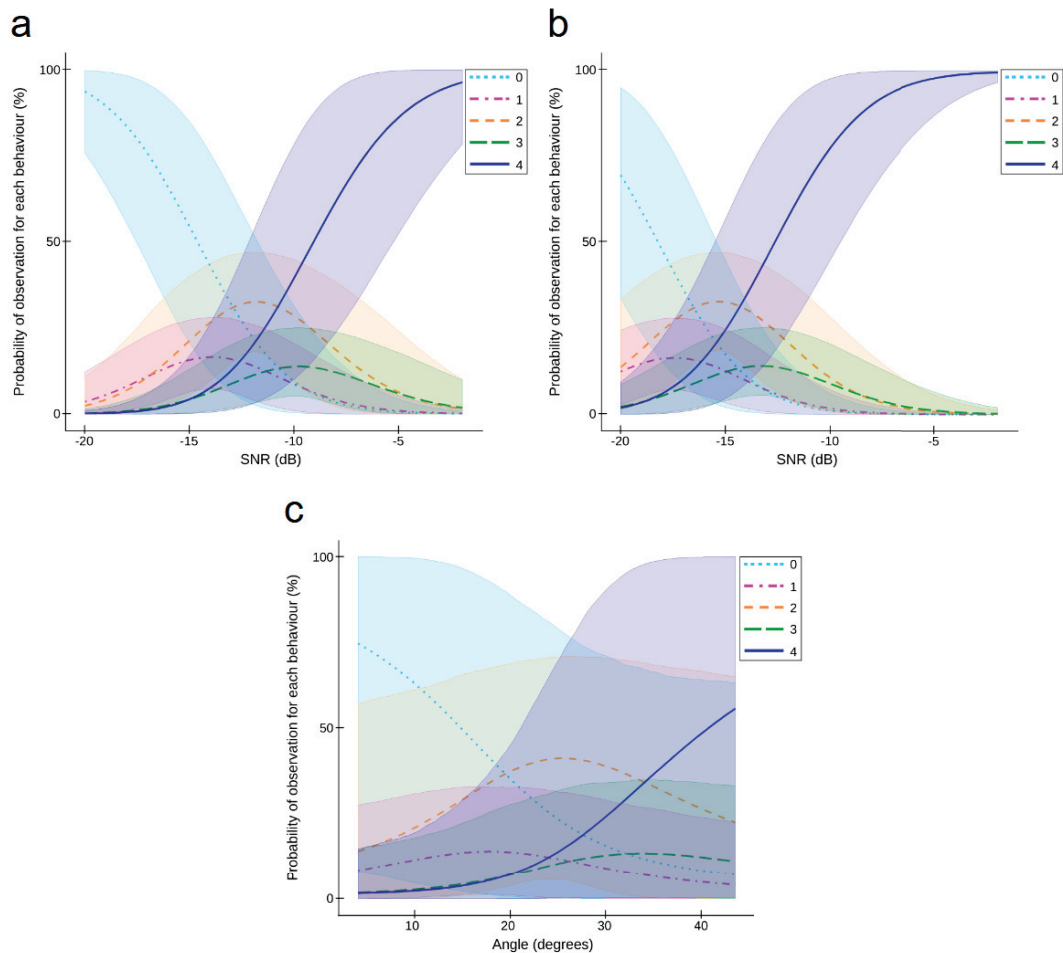


between experimental conditions using a Bayesian approach (see Methods for details). In summary, the probability of behavioral scores was fitted according to two different models: one with the SNR and the position of target source (i.e. either co-located or separated) as fixed factors, and another one with only the data obtained in the separated condition, with the SNR and the initial angle of separation between the target and the masker as fixed factors.

The playbacks revealed that the female's response depended on the SNR between the target and the masker, with higher SNRs inducing higher behavioral scores (Bayesian ordinal model:  $\beta_{\text{SNR}} = 0.32$ , 95% CI = [0.21, 0.45], probit scale; see appendix Supplementary Figures III.2). They also provided strong evidence of an effect of the separation of the target source from the masking noise source, with higher behavioral scores being more likely in the separated condition compared to the co-located condition ( $\beta_{\text{separation}} = 1.15$ , 95% CI = [0.63, 1.69]; appendix Supplementary Figure III.2). By fitting the probabilities of the behavioral scores in function respectively of the SNR and of the co-located and separated conditions, we confirmed that the females' motivation to move towards the target loudspeaker depended highly on the SNR, with lower SNR levels eliciting a female reaction in the separated condition (Figures III.2a and III.2b). Accordingly, the separated condition decreased both the SNR threshold from which the females began to respond and the SNR threshold eliciting a full response.

The signal detection threshold (SNR value corresponding to a 50% probability of a behavioral score equal or higher than 1) was -14.6 dB in the co-located condition while it dropped to -18.2 dB in the separated condition. The full response threshold (SNR value corresponding to a 50% probability of a behavioral score of 4) was equal to -9.1 dB in the co-located condition and -12.7 dB in the separated condition. Both thresholds (i.e. signal detection threshold and full response threshold) lead to a SRM effect equal to 3.6 dB. Interestingly, we found neither an effect of the SNR ( $\beta_{\text{SNR}} = -0.06$ , 95% CI = [-0.41, 0.27], skew normal distribution; appendix Supplementary Figure III.5a) nor of the relative positions of the target and noise sources relative positions ( $\beta_{\text{separation}} = -0.88$ , 95% CI = [-3.93, 2.44]; appendix Supplementary Figure III.5b) on the females' reaction time once the stimulus is detected. Thus, while SRM helps the animal to detect a signal in a noisy environment, it does not seem to influence the delay between the detection and the behavioral reaction.

We then tested for an effect of the angle of separation between the target and the noise sources on the females' responses by focusing only on the separated condition (target loudspeaker separated from the masker; min angle = 4°, max angle = 44°). Figure III.2c shows the fitted



**Figure III.2. Effect of the Signal-to-Noise ratio (SNR) and of the loudspeakers' spacing on the behavioral reaction of female caimans to the playback of young distress calls (Experiment 1, Pantanal, field conditions with wild animals).** **a** Effect of the SNR on the females' response to sound stimuli when the target and the noise loudspeaker are at the same location ("co-located" condition). The probability of eliciting a higher behavioral response increases with SNR (fitted probabilities of behavioral scores: mean of posterior distribution and 95% credible intervals). **b** Effect of the Signal-to-Noise ratio (SNR) on the females' response when the target and the noise loudspeaker are spaced by a minimum angle of  $4^\circ$  ("separated" condition, mean angle between loudspeakers =  $18^\circ$ , min-max =  $4-44^\circ$ ). The females' behavioral reactions are elicited by stimuli with lower SNR compared to the "co-located" condition, supporting the hypothesis that the tested females perform Spatial Release from Masking. **c** Effect of the speaker spacing on the females' response in the "separated" condition. The probability of the females approaching the loudspeaker increases as the separation angle between the target and the noise loudspeakers increases.





probabilities of each 0-4 behavioral score as a function of the angle of separation, while controlling for the SNR. The results support the hypothesis that the larger the angle, the stronger the female's response ( $\beta_{\text{angle}} = 0.10$ , 95% CI = [-0.01, 0.23] on the probit scale; 95.9% confidence that higher angles of separation between the target and the masker elicited higher behavioral scores; appendix Supplementary Figure III.1a).

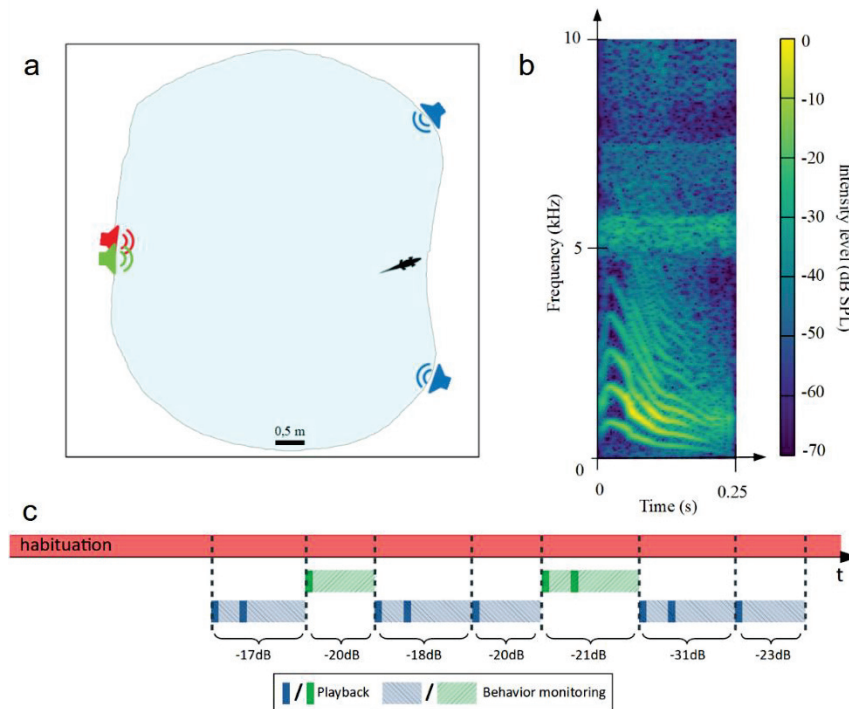
## 2.2 Spatial Release from Masking by juvenile crocodilians

This second experiment was performed on young naive Nile crocodiles *Crocodylus niloticus* in captivity (3-months old juveniles, N = 8). We tested their ability to detect a target signal against noise by playing back series of "contact" calls in a noisy environment (see Figure III.3b for the spectrogram of one call). For each experiment, a crocodile was placed in a large outdoor pool (diameter 8 meters) where a loudspeaker placed on the edge of the pool was continuously emitting a broadband noise.

Several hours later during the night, we played back series of target signals from other loudspeakers placed at different locations around the pool (one "co-located" loudspeaker side-by-side to the noise loudspeaker and two "separated" loudspeakers, Figure III.3a; when non-null, the angle between the separated loudspeakers and the noise loudspeaker varied between 44° and 156°, see raw data in appendix Supplementary Figure III.1b). Each subject was challenged several times with an interval of at least 10 minutes between trials (Figure III.3c; 7-11 trials per subject; total of 30 "co-located" and 41 "separated" trials; see appendix Supplementary Table III.2 for details of the signals played back to each juvenile). For each trial, we assessed the crocodile's ability to detect the target signal against the background noise by rating its behavior according to a binary scale: no orientation towards the loudspeaker emitting the target calls = score 0; orientation towards the loudspeaker = score 1. For the purpose of analysis, we further modelled this scoring using a Bayesian logistic regression (Bernoulli distribution).

The playback tests showed that the ability of the juvenile crocodiles to detect the target signal against the background noise depended both on the signal-to-noise ratio, with higher SNRs inducing higher probabilities of detection ( $\beta_{\text{SNR}} = 0.23$ , 95% CI = [0.10, 0.39], logit scale; appendix Supplementary Figures III.3a), and on the source position, with a higher detection probability when the noise and the target loudspeakers were spatially separated ( $\beta_{\text{separation}} = 1.57$ ,

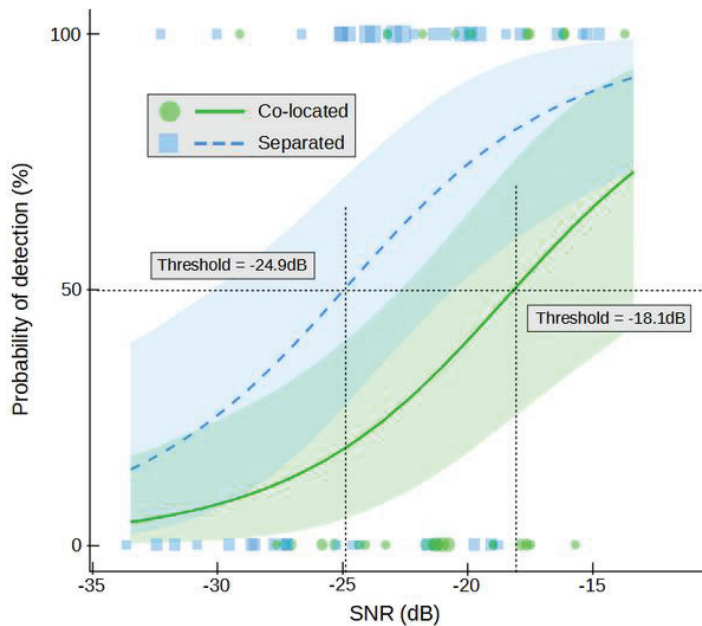
95% CI = [0.40, 2.90]; appendix Supplementary figure III.3b). These results are in line with those obtained in the field experiments reported in the previous section of the article.



**Figure III.3. Experiments on young Nile crocodiles in captivity (Experiment 2).** **a** Schematic representation of the experimental design. A noise was continuously emitted by the "noise" loudspeaker (red). The stimuli were emitted either by the "co-located" loudspeaker (green) or one of the "separated" loudspeakers (blue). **b** Spectrogram of a contact call from a young Nile crocodile. **c** Timeline of an experiment. The masking noise is emitted continuously (red solid line). The target signals (with different SNR) are emitted either by the co-located loudspeaker (green solid line) or by one of the separated loudspeakers (blue solid line). The same signal could be played again by the same loudspeaker if the crocodile had not moved 90 seconds after the end of the first emission. The light dashed lines represent the behavior recording following the target emission.

By modeling the signal detection probability in function respectively of the SNR and of the co-located and separated conditions (Figure III.4), we found that the signal detection threshold (SNR value corresponding to a 50% probability of signal detection; for comparison purpose, this would correspond to a score equal or above 3 in the first experiment) was -18.1 dB in the co-located condition while it decreased to -24.9 dB in the separated condition (i.e. SRM amount equal to 6.8 dB). In accordance with this result, a separated target had a 65.4 % probability of being detected for an SNR of -22.1 dB (median value) while this probability was only 28.6 % for a co-located target (95 % CI = [8.9, 59.8]).





**Figure III.4. Effect of the Signal-to-Noise ratio (SNR) and of the loudspeakers spacing on the behavioral reaction of young Nile crocodiles to the playback of contact calls** (Experiment 2, Crocoparc Zoo, freely moving animals in a large basin; curves = fitted probabilities of behavioral scores: mean of posterior distribution and 95% credible intervals; green dots represent individual trials in co-located condition, blue squares are individual trials in separated condition).

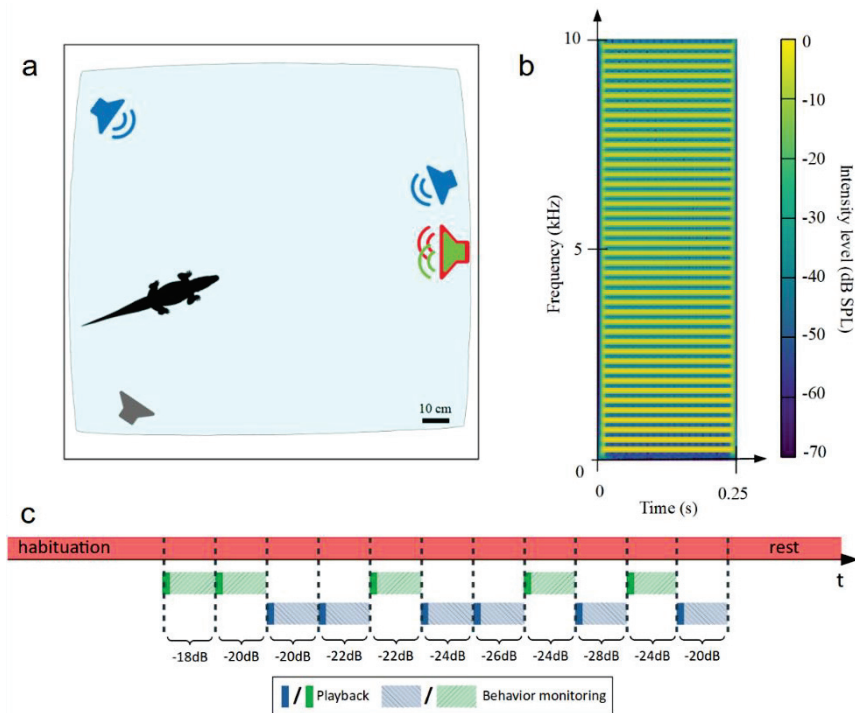
The probability of signal detection increases with SNR in both conditions, i.e. when the target and the noise loudspeaker are close together ("co-located" condition) or spaced apart ("separated" condition). The crocodiles' behavioral reactions are elicited by stimuli with lower SNR in the "separated" condition, supporting the hypothesis that tested young Nile crocodiles perform Spatial Release from Masking. The difference between both detection thresholds (amount of spatial release) is 6.8 dB.

As for the field experiments, we observed no influence of the SNR or of the position of the target loudspeaker on the reaction time ( $\beta_{\text{SNR}} = -1.31$ , 95% CI = [-3.42, 0.78];  $\beta_{\text{separation}} = -11.44$ , 95% CI = [-32.28, 8.74]; appendix Supplementary Figure III.5c and III.5d). We further tested whether increasing the angle between the noise and the target loudspeaker from 44° (minimum angle in the separated condition) to 156° (maximal angle) could improve the crocodiles' ability to detect the target signal and found no effect ( $\beta_{\text{angle}} = -0.01$ , 95% CI = [-0.04, 0.03] on the logit scale; appendix Supplementary Figure III.3c).

### 2.3 Spatial Release from Masking by crocodilians to detect a non-biological signal

This third experiment was performed in laboratory conditions with two juvenile Nile crocodiles (3 years-old). Prior to the experimental procedure, both crocodiles were trained with a Go/No-Go procedure to swim towards a target loudspeaker emitting a synthesized harmonic complex tone (buzz, Figure III.5b; see Methods). The crocodiles' ability to detect the target loudspeaker

against a background noise was then tested during several experimental sessions (65 and 55 trials with crocodile 1 and 2, respectively; see appendix Supplementary Table III.3 for details).



**Figure III.5. Go/No-Go experiments on juvenile Nile crocodiles in captivity (Experiment 3).**

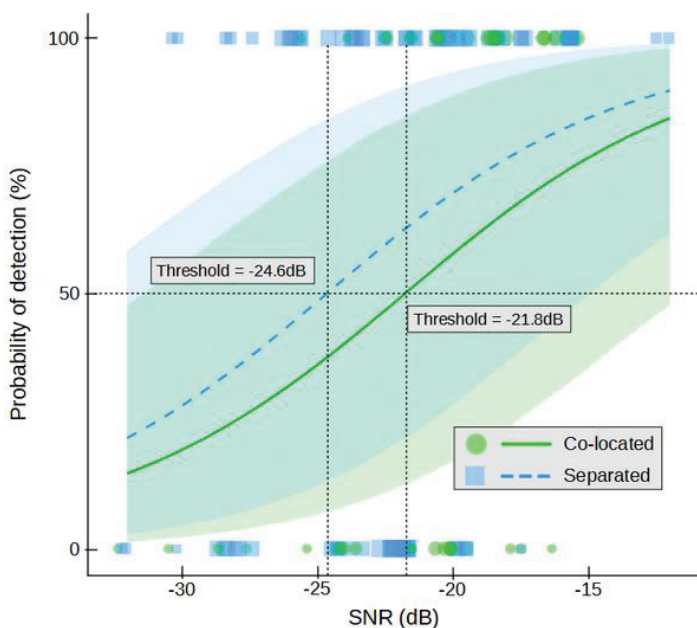
**a** Schematic representation of the experimental design. A noise was continuously emitted by the "noise" loudspeaker (red). The stimuli were emitted either by the "co-located" loudspeaker (green) or one of the "separated" loudspeakers (blue). **b** Spectrogram of the synthetic buzz used as the sound stimulus. **c** Timeline of an experiment. The masking noise is emitted continuously (red solid line). The target signals (with different SNR) are emitted either by the co-located loudspeaker (green solid line) or by one of the separated loudspeakers (blue solid line). The same signal could be played again from the same loudspeaker if the crocodile had not moved 45 seconds after the end of the first emission. The light dashed lines represent the behavior recording following the target emission.

For each session, one of the crocodiles was placed in an experimental pool in a sound-proofed chamber (Figure III.5a). One loudspeaker was continuously emitting white noise. The target signal (sequences of three synthetic signals identical to the ones used during training) was emitted either by the noise loudspeaker (the noise and the target signals were mixed) or by one of two other loudspeakers placed at other locations on the edge of the pool (Figure III.5c). The tests were done in complete darkness. For each trial, we assessed the crocodile's ability to detect the target signal against the background noise by rating its behavior according to a binary scale: no orientation towards the loudspeaker emitting the target calls = score 0; orientation towards



the loudspeaker = score 1. For analysis purpose, we further modelled this scoring using a Bayesian logistic regression.

The playback tests showed that the SNR of the target stimuli had a strong effect on the crocodiles' ability to detect the signal against the background noise, with higher SNRs inducing higher probabilities of detection ( $\beta_{\text{SNR}} = 0.20$ , 95% CI = [0.10, 0.32], logit scale; appendix Supplementary Figures III.1c for raw data and III.4a). Although the effect of the target loudspeaker location (co-located versus separated) appeared weaker than in the two previous experiments, there was a 90.7% probability that separated targets were better detected than co-located ones ( $\beta_{\text{separation}} = 0.58$ , 95% CI = [-0.28, 1.47]; appendix Supplementary Figure III.4b). By modelling the signal detection probability according to, respectively, the SNR and of the co-located and separated conditions (Figure III.6), we found that the signal detection threshold (SNR value corresponding to a 50% probability of signal detection) was -21.8 dB in the co-located condition while it decreased to -24.6 dB in the separated condition (i.e. spatial release from masking equal to 2.8 dB). A separated target had a 63% probability of being detected for an SNR of -22 dB (median value) while this probability was 49% for a co-located target (95% CI = [-5.70, 32.74]).



**Figure III.6. Effect of the Signal-to-Noise ratio (SNR) and of the loudspeakers' spacing ("co-located" versus "separated") on the behavioral reaction of juvenile Nile crocodiles to the playback of a synthetic buzz** (Experiment 3, ENES Laboratory; the animals have been trained to move towards the target loudspeaker; curves = fitted probabilities of signal detection: mean of posterior distribution and 95% credible intervals; green dots represent

trials in co-located condition, blue squares are individual trials in separated condition). The probability of target signal detection increases with SNR in both "co-located" and "separated" conditions. The tested crocodiles detect stimuli with lower SNR in the "separated" condition, supporting the hypothesis that they perform Spatial Release from Masking. The difference between both detection thresholds (amount of spatial release) is 2.8 dB.

As in the two other experiments, we found no effect of the SNR or of the angle between the two loudspeakers on the crocodile's reaction time ( $\beta_{\text{SNR}} = 0.05$ , 95% CI = [-0.43, 0.58];  $\beta_{\text{separation}} = 2.06$ , 95% CI = [-2.13, 6.72]; appendix Supplementary Figure III.5e and III.5f). We further tested whether increasing the angle between the noise and the target loudspeaker from 16° (minimum angle in the separated condition) to 178° (maximal angle) could improve the crocodiles' ability to detect the target signal and found no effect ( $\beta_{\text{Angle}} = 0.00$ , 95% CI = [-0.01, 0.01]; appendix Supplementary Figure III.4c).

### 3. Discussion

In this study, we tested whether crocodylians use Spatial Release from Masking to detect a sound target against a continuous background noise. We performed experiments in three different contexts: in the field with wild adult animals, in a naturalistic setup in captivity with naive juvenile subjects, and in a laboratory Go/No-Go experiment with trained juvenile subjects. The combined results of these three experimental approaches confirm that crocodylians detect sound signals better when the target source is spatially separated from the masking noise source, suggesting that these animals use SRM in their daily lives.

Conducting experiments with crocodiles can be challenging. In the field and in the zoo, they habituate to played back signals very quickly, which limits the number of trials performed with a given individual. In the field, to ensure as much as possible that each female could be tested in both co-located and separated conditions for several SNR, we choose to present successively the signals starting from the lowest SNR until it elicits a response from the animal. To avoid a potential cumulative effect due to this protocol, we took several precautions: (1) we were very careful to note the smallest observable behavioural response suggesting a possible detection of the signal, (2) we repeated successively the same signal 3 times to allow the female the opportunity to respond when they hesitated, and (3) we leave a significant temporal delay between the stimuli. In the zoo as in Go/No-Go experiments in the laboratory, we optimized the number of trials by placing several speakers around the ponds in order to change the origin of the sound. Go/No-Go experiments in the laboratory required extensive training of the animals. This time-consuming training, combined with the logistical constraints inherent in these animals when kept in captivity, also limits the number of subjects that can be included in the experiments. In addition, the ectothermy of these animals imposes a long delay between experimental sessions for the animal to regain hunger and be sufficiently motivated to perform



the task. These constraints explain why there are some gaps in our data, both in the range of SNRs tested and in the range of separation angles (appendix Supplementary Figure III.1). Such limitations impact the statistical power of the analyses and call for caution in interpreting results. Nonetheless, our data highlight that the spatial separation between the target and noise sources has a major influence on the detectability of the target source. While an increase in SNR improves the signal detection ability of the tested individuals in both types of experimental conditions ("co-located" and "separated"), detection thresholds are always lower when the target and noise sources are spatially separated.

Because the three sets of experiments (field, zoo, and Go/No-Go) differ in terms of speakers' position, distance between speakers and tested individuals, and, most importantly, in terms of biological context and crocodilian species, the absolute values of detection and response thresholds cannot be accurately compared. The SRM size effect, however, remains close in magnitude (3.6 dB, 6.8 dB, and 2.8 dB, in the field, in captivity, and in Go/No-Go experiments, respectively), and can be considered representative of crocodilian SRM capabilities.

To the best of our knowledge, this study is the first to investigate SRM in wild animals freely behaving in their natural habitat, and to combine this approach with investigations in captivity and in the laboratory. Both in the field and in the zoo, we did not train the animals to respond. Consequently, the behavioral reaction of the tested subjects to the stimuli was likely modulated by several factors influencing their internal motivation. Therefore, we may have underestimated the ability of animals to detect the target signal, and the amplitude of SRM may be greater than reported. In the field for instance, female Yacare caimans remained close to their own young when challenged with the target signals. Moving toward the target loudspeaker meant that the females had to abandon their young. This trade-off between motivation to stay and motivation to go may have decreased the females' reaction level. It probably explains some of the variation between individual responses. To understand this variation, it would have been interesting to know the number of nestlings present with each tested female, and to monitor the vocal activity of the young as both may have influenced the mother's decision. In the zoo context, juvenile Nile crocodiles who found themselves isolated for the duration of the experiment, may have faced another type of behavioral trade-off between swimming to a speaker mimicking a sibling and remaining still to limit predation risk, as young crocodiles are heavily predated in the wild. Conversely, in the Go/No-Go laboratory experiment, the tested subjects have been trained to move toward the loudspeaker by getting a food reward. Thus, it is likely that the subjects' motivation to respond to the target stimuli was high, and at least, fairly consistent over the

course of the experimental trials. Nonetheless, we still observed variability in the animals' response, potentially related to personality differences and also probably to their bradymetabolism differences punctually affecting their motivation to perform the experiment.

Contrary to our expectations, the SRM values obtained with the Go/No-Go procedure were lower than in the two other contexts. One possible explanation lies in the acoustic environment in which the experimental trials were performed. The test booth was quiet (background level = 40 dB SPL), but not perfectly anechoic (reverberation time = 0.44s, volume of the booth = 9.11m<sup>3</sup>). The tested crocodiles may have perceived some early acoustic reflections in addition to the direct sound waves which may have decreased the ability to detect the target signal.

A second possible explanation for this lower SRM value could come from the nature of the target signal used in this experiment. In humans, the SRM, also related to the cocktail party effect, involves energetic aspects (i.e. energetic masking) as well as cognitive aspects (i.e. informational masking (Freyman et al., 2001)). In the first two experiments, the target signals were biologically relevant to the crocodile, unlike in the last experiment (synthetic buzz). This may have modulated the amount of informational masking across experiments and contributed to the weaker SRM effect in the Go/No-Go experiment.

In humans, the mechanisms underlying SRM have been extensively explored and reviewed (Bronkhorst, 2000; Jones & Litovsky, 2011; Litovsky, 2012). First, when the target and masker are spatially separated, half of SRM effect comes from the “better ear effect”, where the SNR is more favorable in one ear (due to noise attenuation by the head shadow) than in the other. This effect is purely monaural. Second, the ability of the auditory system to utilize binaural aspects of the signal, including time (ITD) and level (ILD) differences between the ears is also known to contribute to SRM. Third, “binaural summation” (i.e. the fact that a signal presented to the front will activate both ears and then make that sound easier to hear due to the summation of the signals at both ears) provides an additional contribution to SRM. In our study, a reliable SRM effect was reported in all experiments, and all three mechanisms described in human could also have contributed to the observed SRM effect. In particular, it is now well known that crocodile ears are acoustically coupled by air-filled cranial sinuses (Bierman et al., 2014; Carr et al., 2016), which greatly increases directional cues such as ITDs (Carr & Christensen-Dalsgaard, 2015). At the encoding level, alligators have been shown to form ITD maps in the brainstem nucleus laminaris similarly to birds, again suggesting a convergence among modern archosaurs (Kettler & Carr, 2019). The crocodylian binaural system may therefore be as well-





developed as that of birds and thus could be effective in detecting spatially separate signals (Bierman & Carr, 2015). However, the size of each effect might have been overestimated or underestimated because of the fluctuating position of the crocodile head during stimulation. In fact, under some conditions, the masker and separated target could be played on the same side of the crocodile's head, altering the magnitude of the better-ear-effect and/or the magnitude of the summation effect. Therefore, tested in freely-moving animals, the potential SRM effect may not have been maximized in all trials. In conclusion, on the one hand, the relative contributions of the monaural better-ear-effect, binaural cues and binaural summation effect for SRM remain unknown for non-human animals, including crocodiles, and would require further study. On the other hand, our study supports a global SRM effect in the field, regardless of head position relative to the source and target positions.

Despite these differences between the three experimental conditions, our results highlight the importance of loudspeaker spacing for each of the three. Strikingly, the field experiment shows that the SRM increases significantly when the separation angle between the target and the noise loudspeakers increases from  $4^\circ$  to  $44^\circ$ . Since we did not find this angle effect in Experiment 2 (angles ranging from  $44^\circ$  to  $156^\circ$ ) or Experiment 3 (angles ranging from  $16^\circ$  to  $178^\circ$ ), we assume that the SRM quickly becomes effective at small angles. Interestingly, Papet et al. (2020) reports that the minimum audible angle (MAA) is about  $13.3^\circ$  in crocodilians. This threshold suggests that the effect of angle on SRM may be dominant for small angles in the range  $4^\circ$ - $15^\circ$ , and becomes saturated for higher angular values.

Our results still support the hypothesis that SRM is a shared ability among vertebrates. Gray treefrogs showed SRM ranging from 3 to 12 dB (Bee, 2007; Nityananda & Bee, 2012). Despite the enormous variability in SRM as a function of experimental context (Bee & Micheyl, 2008), birds develop high abilities to use spatial cues as a means to detect a target signal. For example, budgerigars *Melopsittacus undulatus* display a SRM of around 9 dB when required to detect pure tones against white noise in a Go/No-Go experimental setup (Dent, 1997), but achieve an impressive SRM of 20 to 30 dB in an identification task with biological signals (Dent et al., 2009). In mammals, the SRM reaches 10 dB in ferrets (Hine et al., 1994), and 12 to 19 dB in pinnipeds (Holt & Schusterman, 2007). In humans, the SRM has been estimated to be between 15 and 18 dB with "clicks" as target signals, the masker being broadband noise (Sabeti et al., 1991). These high values could be explained by a greater ability to analyze auditory sound scenes, by different experimental conditions, or, in humans, simply by the fact that subjects are better able to understand the task required for the experiment. The amount of SRM measured

in ethological studies is likely to be lower than in neurophysiological studies, due to perceptual and decision-making effects. This makes it difficult to compare values found by an ethological approach such as the one employed here with values measured with a neurophysiological approach.

In conclusion, our several approaches - from field to laboratory experiments - demonstrate the use of SRM in crocodylians, and highlight that these amphibious animals can take advantage of the spatialization of sound sources in their natural environment to analyze sound scenes, and improve detection of signals containing relevant information. We argue that naturalistic approaches are absolutely necessary to fully understand and measure SRM abilities. For this and other biological processes, field experiments provide the ultimate proof of the relevance of a mechanism. The difficulty of conducting such experiments is offset by the naturalistic results they provide.

## 4. Methods

Supplementary figures and tables are presented in the appendices at the end of the manuscript. Dataset, codes, videos and audio signals supporting the present results can be found in the Zenodo repository (Thévenet et al., 2022a): <https://doi.org/10.5281/zenodo.5971364>.

### 4.1 Experiment 1 (Field experiment): Spatial release from masking during mother-young communication in wild

#### *Field location and tested animals*

We conducted the field work at "Nhumirim ranch" (Embrapa Research Station, Mato Grosso do Sul, Brazil; 18°59'16.1"S 56°37'08.8"W), an area that covers 4310 ha with about 100 lakes (Campos & Magnusson, 1995; Campos et al., 2015). We first surveyed the area for nests and Yacare caiman females in February - March 2019, and then conducted the playback experiments at the end of the hatching season (April 30th - May 11th 2019). We tested 16 adult females that had been previously identified as having built a nest and laid eggs. Most of the tested females were on separated lakes (10 of 16 individuals, Figure III.1a). When two females living in the same lake were tested successively (3 lakes x 2 individuals = 6 individuals) we always chose individuals separated by at least 100 meters, and carefully checked that the second



female to be tested could not have heard the sound stimuli broadcast to the first tested female. To avoid habituation, each female was involved in only one experimental session. All experiments were conducted during the day.

### ***Experimental signals***

We tested females with juvenile distress calls (Figure III.1c), which are well-known to elicit protective behavior from the mother (Vergne et al., 2009). The day before the first experimental session, we recorded distress calls from 3 juveniles approximately 3 weeks old. Calls were elicited by successively manipulating each individual. Handling time did not exceed 2-3 minutes and juveniles were immediately returned to their mother after being recorded. These individuals belonged to the same clutch, and their mother was not included in the females tested. Thus, the females tested were all tested with calls from juveniles that were not their own. Previous work has shown that female crocodylians respond indifferently to the calls of their young and the calls of unknown young (Vergne et al., 2007; Vergne et al., 2011).

During the playback experiments, we broadcast a "masking noise" and a "target signal". The masking noise was a white noise (2 hours duration, frequency range [20, 20000] Hz; 83 unweighted dB SPL measured at 1m using a Sound level meter AMPROBE SM-10; slow time window equal to 1 second). It was broadcast in a loop for the duration of each experimental session. The target signals were designed as sequences of 10 successive distress calls (randomly selected from our bank of recorded calls). Each call was previously low-pass and high-pass filtered (cut-off frequencies: 20 Hz and 10 kHz respectively, 3rd order filters), and normalized in intensity by its RMS value (i.e. each call contained the same amount of energy). In each target signal, the duration of silence between two calls varied randomly between  $1.25 \pm 0.25$  s to reproduce a natural rhythm (total duration of the target signal = 17 s). We created 11 target signals, which differ from each other in their sound level. The intensity of the calls within each target signal was precisely adjusted to the intensity of the masking noise in the range [-20, 0] dB with a step size of 2 dB. The signal-to-noise ratios between the target signals and the masker were then computed directly from the intensities of the audio signals.

### ***Playback protocol***

Prior to an experiment, we placed three remote-controlled loudspeakers (FoxPro Fusion, rear loudspeaker, see appendix Supplementary Figures III.6a and b for the technical specifications) just above the water surface, approximately 20 meters from the tested female (minimal distance

= 12 m; maximal distance = 50 m; Figure III.1b). Two of the loudspeakers were placed side by side: one played the masking noise ("noise" loudspeaker), and the other was used to play back the target signal ("co-located" loudspeaker). The third loudspeaker ("separated" loudspeaker) was positioned to form an isosceles triangle with the noise loudspeaker and the initial position of the tested caiman female. This equidistance of the speakers from the female allowed us to consider the SNR value at the female's head position as equal to the SNR calculated at the speakers' emission. By estimating the distances between the speakers and the female, we calculated the separation angle  $\theta$  as the angle formed by the female, the noise loudspeaker and the separated loudspeaker. Because the crocodiles were free to move, we could not ensure a constant angle between the female's head and the loudspeaker from trial to trial.

The target signals were alternately emitted from the co-located speaker and the separated speaker. At the beginning of the experiment, the female was at the same distance from the co-located loudspeaker and the separated loudspeaker (Figure III.1b). The masker was played continuously throughout the experimental session, starting with a quick fade-in until it raised to the intensity level of 83 unweighted dB SPL to avoid frightening the female with a sudden noise. We never noticed any change in the females' behavior during the 10 minutes after the masker appeared. Specifically, we did not notice any type of avoidance behavior of the loudspeaker emitting the masker.

Before playing back the first target signal we first observed the female's behavior for at least 5 minutes (Figure III.1d). If the female moved during this observation period, we waited another 5 minutes. If the female's distance from the co-located and separated loudspeakers was no longer equal, we then changed the position of the loudspeakers to recreate the isosceles triangle between the two loudspeakers (Figure III.1b), and we started another 5 minutes observation period before the experiment.

At the end of the observation period, we broadcast the first target signal from the co-located speaker at a low intensity level (SNR varying between -18 and -4 dB). The target signal was emitted 3 times, once per minute (Figure III.1d). However, the delay between these renditions was variable, depending on the female's behavior: if she moved or dived underwater, we waited for her to stop or to reappear at the surface before broadcasting the target signal again. After the third playback of the target signal, we waited at least 3 minutes, then repeated the same procedure this time from the separated loudspeaker. After a post-playback delay of at least 3 minutes, we would emit a new target signal increased by 2 dB, following the same procedure.



An experimental session thus consisted of a repetition of this procedure, alternating the playback between the co-located speaker and the separate speaker, and increasing the signal-to-noise ratio by +2 dB in each cycle. The experimental session was stopped as soon as the female responded to the stimuli by orienting in the direction of the target speaker and/or approaching it. Specifically, we stopped the playback when the female had changed her initial position by more than one body length. In summary, each female was tested with 1 to 9 pairs of target signals (each pair corresponding to a broadcast by the co-located speaker and a broadcast by the separated speaker).

### *Analysis of behavioral reaction to playback*

We observed and filmed the behavior of the females throughout the experiments. Because the field experiments were conducted on wild animals with the ability to express their full range of behaviors, we assessed the response of the females by scoring their behavior as follows (motivation scale): score 0: no behavioral response (no movement); score 1: the female moved her head or body, but not in the direction of the target loudspeaker (misdirected response); score 2: the female moved her head and looked towards the target loudspeaker without moving her body; score 3: the female moved less than 1 body length towards the target loudspeaker; score 4: the female moved more than 1 body length towards the target loudspeaker.

## **4.2 Experiment 2 (experiment in zoo): Spatial release from masking during between-juveniles interactions**

### *Location and animals*

We performed these experiments in October 2019 at the "Crocoparc" zoo (Agadir, Morocco). We worked with naive juvenile Nile crocodiles (*Crocodylus niloticus*) hatched in captivity (n = 8 individuals; three months old;  $36 \pm 2$  cm length). These animals were housed together in an exterior enclosure not visible by the public. They had never been included in any experiments before. Each crocodile subject was tested during only one experimental session.

### *Experimental signals*

As in experiment 1, we broadcast masking noise and target signals. The masking noise (white noise, 2 hours duration) was played continuously in a loop, starting before putting the crocodile in the pond and throughout each experimental session (frequency range [20, 20000] Hz; 83

unweighted dB SPL at 1 meter with the same sound level meter and same settings as in experiment 1). As target signals, we used twelve different sequences of three identical Nile crocodile contact calls from our recording data bank (twelve-unit calls from young Nile crocodile previously recorded in the Okavango Delta, Botswana, by T. Aubin and N. Mathevon; see spectrogram on Figure III.3b). Contact calls are known for maintaining cohesion among juveniles by soliciting their reunification (Vergne et al., 2009). Each call was previously band-passed filtered between 20 Hz and 10 kHz (filter order of 3), and its intensity was normalized by its RMS value. In each target signal, the duration of the silences between the calls was randomly set between  $5 \pm 1.5$  s (to match the natural rhythm), resulting in a total signal duration of 11 seconds. We adjusted the intensity level of the target signals (directly in the audio files, as in Experiment 1) to achieve an SNR in the range [-32, -16] dB with a 2 dB step.

### ***Playback protocol***

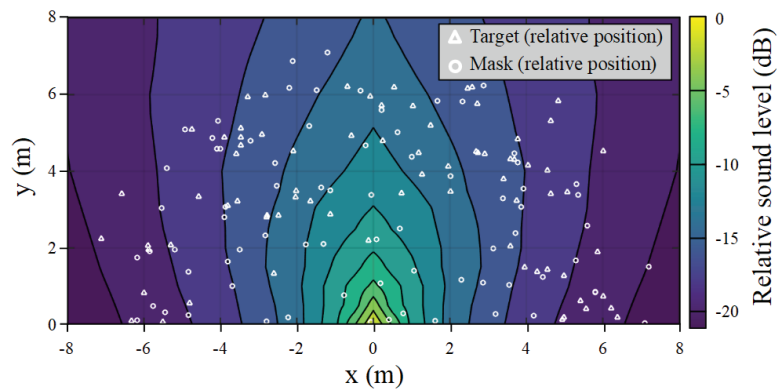
The experiments were performed outdoors at night in an artificial pond of approximately 40 m<sup>2</sup> (maximum dimensions: 6 x 7 meters; Figure III.3a). Four remotely controlled loudspeakers (FoxPro Fusion, rear loudspeaker, appendix Supplementary Figures III.6a and b) were placed on the pond shore (Figure III.3a). As in experiment 1, two loudspeakers were placed side by side: one broadcasting the masking noise (noise loudspeaker) and the other emitting the target signal ("co-located" loudspeaker). The other two loudspeakers were placed at distance from the noise loudspeaker ("separated" loudspeakers). The location of the loudspeakers around the pond was changed between each tested subject to avoid positional bias and to cover a wide range of angles between the target speaker, the noise speaker, and the crocodile's position. Given the size of the pond, the distance between the tested animal and the target loudspeaker was biologically relevant: in the wild, juveniles of the same groups are often one to a few meters apart.

Prior to each trial, the tested juvenile was placed alone in the pond the afternoon before the playback of the target signals (at least 3 hours before dusk), allowing it to become accustomed to its new environment (Figure III.3c). The masking noise was broadcast continuously during this habituation period and throughout the experimental session. During the trials, the experimenters controlled the playback of the target signals while remaining 15m distant from the experimental pond, out of sight of the animal. The first target signal was broadcast by one of the randomly selected target loudspeakers (co-located or separated), at a random SNR value. If the crocodile had not moved 90 seconds after the end of the target signal played, the same



signal was played again on the same loudspeaker. Then, we waited ten minutes after the last signal was played before playing another target signal (randomly chosen from the sound bank) from another target loudspeaker. On average, we performed  $8.8 \pm 1.4$  trials per crocodile tested.

Because the tested crocodile was free to move within the pond, its initial position varied between trials. Therefore, while the SNR value of the played back signal was chosen by the experimenters, the SNR actually perceived by the crocodile at the beginning of the playback depended on its position in the pond relative to the noise speaker and the target speakers. To measure the SNR perceived by the crocodile, we mapped the SNR variations at the pond surface by performing an acoustic propagation experiment. For this propagation experiment, we played back distress calls and noise, and measured their intensity at different points in the pond. This allowed us to model an acoustic map of the pond representing the intensity variations as a function of the position of the crocodile in the pond (Figure III.7). The SNR corresponding to the position of each crocodile tested was then calculated from the intensity levels of the target signal ( $L_T$ ) and masker ( $L_M$ ) and the position of the animal. For each experiment, the initial perceived SNR ( $SNR_p$ ) was defined as follows:  $SNR_p \text{ (dB)} = L_T - L_M$ .



**Figure III.7. Acoustic propagation of a juvenile call and white noise on the pond.** The position of the sound source is normalized at (0, 0). The white triangles and circles represent respectively the positions of the crocodile relatively to the target (triangles) and noise (circles) loudspeakers at the beginning of each experimental trial. The sound intensity level is coded by the color scale.

For each trial, we also measured the separation angle theta formed by the crocodile, the noise loudspeaker and the target loudspeaker. This angle was constrained both by the experimental configuration (Figure III.3a) and by the initial position of the tested juvenile, and varied between  $44$  and  $156^\circ$ . As in Experiment 1, we were unable to ensure a constant angle between the crocodile's head and the masker and/or the target from trial to trial.

### *Analysis of behavioral reaction to playback*

We observed and filmed the behavior of the juveniles during all the trials (infrared cameras ABUS TVCC34010). The videos were analyzed using Kinovea software ([www.kinovea.org](http://www.kinovea.org)). In order to accurately measure the positions in the field and the distances travelled by the crocodiles, we took care to correct the distortion of the camera lens and the geometric perspective error. We extracted the position coordinates of the loudspeakers and the crocodile (the point between the eyes) at the beginning of each playback. Based on these coordinates, we calculated the separation angle  $\theta$  between the crocodile, the noise loudspeaker, and the target loudspeaker.

To assess the juvenile's response to the target signal, we used a binary scale (detection scale), giving a score of 1 if the juvenile showed significant orientation or movement toward the target loudspeaker and 0 if it still had not responded at the end of the playback of the target signal. We also measured the reaction time (in seconds) between the first observed behavioral response and the stimulus onset.

## **4.3 Experiment 3 (Go/No-Go experiment in the laboratory): Spatial release from masking with trained animals**

### *Location and animals*

We conducted these experiments between March and June 2019 at the ENES laboratory. We worked with two Nile crocodiles born in captivity at the zoo "La Ferme aux Crocodiles" (Pierrelatte, France). These animals were three years old and they were housed at the ENES animal facilities. They had been previously included in an experiment on sound localization (Papet et al., 2020), involving a conditioning procedure using harmonic complex tones (buzz). In this experiment, each subject was tested once a week for 14 weeks.

### *Experimental signals*

As in Experiments 1 and 2, we broadcast a masking noise and target signals. The masking noise (white noise, 2 hours duration) was played continuously in a loop before the tested subject was placed in the experimental room and throughout each trial (frequency range [20, 20000] Hz; 60 dB SPL at 50 cm). The target signals were sequences of three different synthetic buzzes (harmonic complex tones; fundamental frequency  $f_0 = 208, 220$  and  $233$  Hz; duration = 500 ms





each; signals synthesized with Python 3.7, SciPy; Figure III.5b). Each target signal was designed as a repetition of three identical buzzes, separated by an interval of 2 seconds  $\pm$  500 ms (total duration of each target signal = 9 seconds). The intensity level of the target audio signals was adjusted to achieve a signal-to-noise ratio in the range [-32, -16] dB with a 2 dB step, as in experiments 1 and 2.

### ***Behavioral conditioning***

Prior to the experiment, the two Nile Crocodiles were trained twice a week to come towards a sound source. The training followed a classical Go/No-Go procedure (Papet et al., 2020): two speakers were placed at the edge of the pool, with only one emitting target signals. As soon as the crocodile touched the target speaker with its snout, it was rewarded with a piece of meat. Before and after the test period, both individuals achieved 100 % success in the conditioning sessions.

### ***Playback protocol***

The experiments were conducted in the dark in a dedicated sound attenuation chamber (TipTopWood©, dimensions = 1.8 x 2.3 x 2.2 meters, background noise < 40 dB SPL, reverberation time = 0.44 s; Figure III.5a), where a squared pool (1.75 m wide) had been set up for the purpose of the experiment. The pool was filled with water to a level that allowed the crocodile to swim (water depth = 10 cm; Papet et al., 2019). Four loudspeakers (AudioPro, Bravo Allroom Sat, appendix Supplementary Figures III.6c and d) were installed just beyond the water surface at the edge of the pool. During each trial, a loudspeaker continuously broadcast the masking noise ("noise" loudspeaker). In the co-located condition, the same speaker also played the target signal mixed with the noise. For the separate condition, two speakers placed at different locations could play the target signal. To maintain the motivation of the crocodiles to respond to the target signals, we chose to reward them each time they came to the target speaker during the experimental trials. It was indeed not possible to reinforce the behavioral response of these animals outside of the experiments if we wanted the animals to maintain their motivation to respond to the signals. In front of each speaker was a system that hid food (a small piece of meat) to reward the animal if it approached the target speaker in response to the stimulus. To control for the possible effect of the smell of the food, we placed a fourth speaker, always silent, accompanied like the other three by the system hiding the food

(but which was never delivered to the animal). This loudspeaker was never approached in response to a sound.

With the exception of the co-located/noise loudspeaker, the spatial locations of the loudspeakers were changed between each experimental session. The sound emission chain consisted in two computers and two power amplifiers (Yamaha AX-397) connected to the loudspeakers and placed outside the experimental chamber. We recorded the behavior of the tested subject with an infrared camera (ABUS TVCC34010) connected to a computer. The tested crocodile was released into the pool at least 20 minutes before the start of an experimental session (Figure III.5c). The noise loudspeaker was already on and was not turned off until the end of the experimental session. The first target signal was broadcast either from a separated loudspeaker or from the co-located loudspeaker at a specific intensity level (both parameters were randomly picked). If the crocodile had still not moved 45 seconds after the third buzz of the target signal ended, we repeated the same target signal once. The crocodile was rewarded if it approached the target loudspeaker within 5 minutes of the last buzz. If the crocodile responded correctly (movement toward the target loudspeaker), we waited 5 minutes before starting another trial. On average, we performed  $9 \pm 2$  trials during an experimental session, covering a wide range of SNRs. Each session always included a few trials at high SNR to check the crocodile's motivation to respond. The crocodile was then left 20 minutes in the pool before being recaptured, to limit an association between the final target signal and a stress-inducing event.

We measured the separation angle theta (the angle formed by the crocodile, the noise loudspeaker and the target loudspeaker) at the onset of the playback. This angle was constrained both by the configuration of the experimental set-up (Figure III.5a), and the initial positions of the tested subject, and varied between 16 and 178°. As for the two other experiments, we could not ensure a constant angle between the crocodile's head and the masker and/or target loudspeaker from trial to trial.

### *Analysis of behavioral reaction to playback*

As in Experiments 1 and 2, we observed and filmed the behavior of the tested subjects throughout the experiments. Video analyses were performed using Kinovea software. Before video analysis, we corrected for camera lens distortion and geometric perspective error. We measured the position coordinates of the loudspeakers and the initial positions of the crocodile (using the point between the eyes) before the start of the playback. Based on these coordinates,



we calculated the separation angle  $\theta$  between the noise loudspeaker, the crocodile and the target loudspeaker.

To assess the tested subject's ability to detect the target sound against the background noise, we used the same binary scale as in the second experiment (detection score), giving a score of 1 if the juvenile showed orientation or movement toward the target loudspeaker, and 0 if it still did not respond within the 5 min observation period following the last buzz of the target signal. If the crocodile did not respond more than twice to one of the higher SNRs (-16 dB or -18 dB) in the same session, the entire session was excluded from the final data set, considering that the motivation to respond to the target signal was not sufficient (only one session had to be excluded). We also measured the latency to respond, i.e. the time between the animal's first response and the preceding target signal.

## 4.4 Statistics and Reproducibility

All statistical analyses were performed in R (v.3.6.2) from a Bayesian perspective, which provides more flexible and considerably richer investigations than the frequentist approach. Bayesian algorithms also have the advantage to be robust for any sample size. Each of the behavioral responses was investigated using this approach: the behavioral response ([0-4]; Experiment 1) score or the signal detection (0 or 1; Experiments 2 and 3), and the latency time to react. These variables were modelled using Bayesian mixed models with random intercepts per tested crocodiles, fit in Stan computational framework accessed with brms package (Bürkner, 2017).

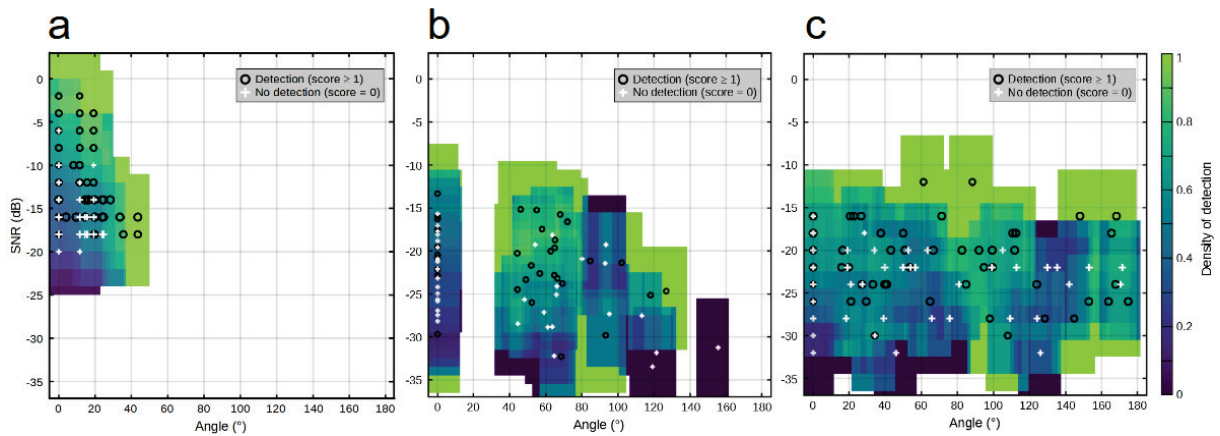
Behavioral scores were modeled with a cumulative link function, a powerful model that is too often underestimated and left out for ratings data (ordinal regression) (Bürkner & Vuorre, 2018). The most appropriate link function was chosen by selecting the most predictive models. Two independent cumulative models were constructed as follows: a first one with the SNR and the position of target source (i.e. either co-located or separated) as fixed factors, and a second one by focusing only on the data obtained in the separated condition, with the SNR and the initial angle of separation between the target and the masker as fixed factors. Detection scores were modeled using the Bernoulli distribution (logistic regression), with SNR and position of the target source (co-located or separated) as fixed factors. This first model allowed us to approximate the signal detection threshold T50 (corresponding to a 50% probability of target detection (Bee & Schwartz, 2009) in both the co-located and the separated condition. As in the

first experiment, a second model was based only on the data obtained in the separated condition and included the SNR and the initial angle of separation as fixed factors.

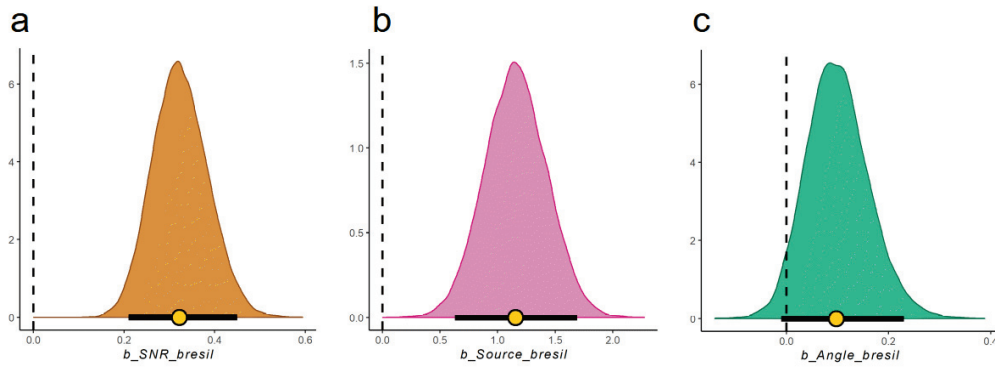
Finally, when signals were detected, the effect of SNR and the position of the target source on the animals' reaction time was investigated using a skewed normal distribution to consider its asymmetry. All models were based on four chains of 10000 iterations with 2000 warmup samples. Model convergence was checked with traceplots and the Gelman-Rubin's potential scale reduction factor (Rhat equal to 1.00; Gelman et al., 1992) on split chains. The interaction between fixed factors was tested and removed from each model, as it reduced the fit of the models (WAIC calculated based on the posterior likelihood; Gelman et al., 2014). The regression coefficients for each model were summarized using the mean of their posterior distribution and the 95% credible interval, reported in the text as 95% CI. Contrasts were reported using the median of the posterior distributions and the 95% credible intervals.

## 5. Supplementary information

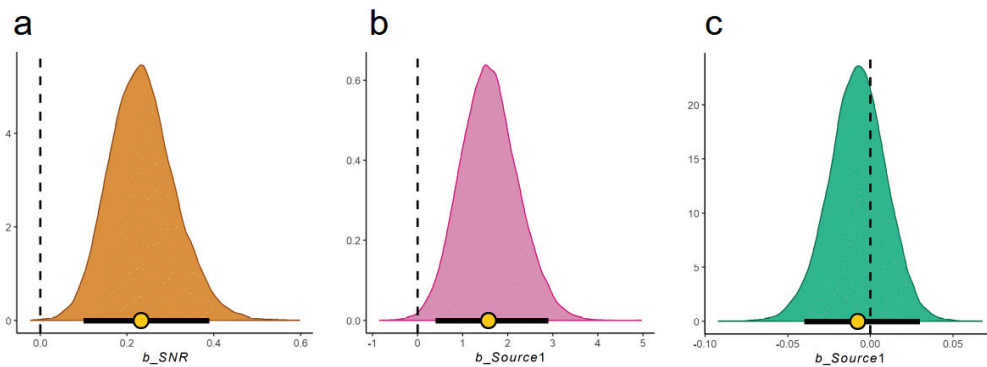
### 5.1 Supplementary Figures



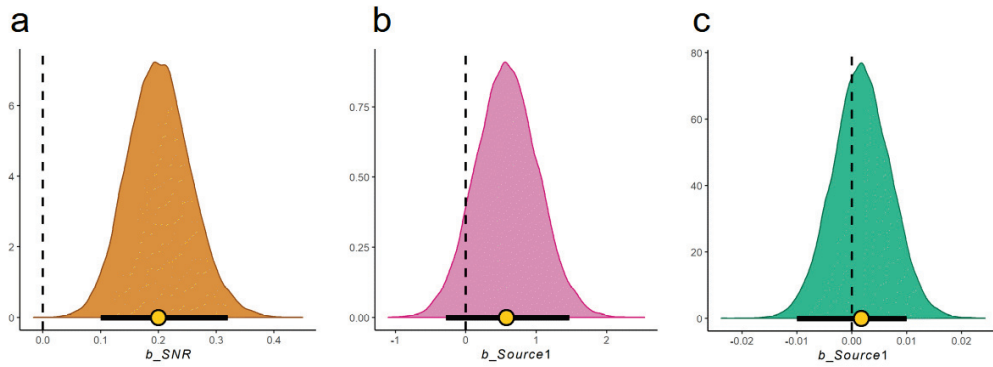
**Supplementary Figure III.1. Signal detection matrix.** Each performed trials are symbolized either by a “o” (no reaction) or a “+” (at least a head or a body movement towards the target loudspeaker). The probability of detection is emphasized by the green-blue color scale, ranging from 0 (deep blue, no detection) to 1 (light green, effective detection; probabilities calculated using a sliding squared window  $24^\circ \times 12$  dB). **a** Ability of female caimans to detect a sound stimulus (young distress call) as a function of Signal-to-Noise Ratio (SNR) and loudspeakers spacing (104 trials on 16 females). The intensity of the behavioral reaction is not reported here. **b** Ability of young Nile crocodiles to detect a sound stimulus (contact call) as a function of Signal-to-Noise Ratio (SNR) and loudspeakers spacing (71 trials on 8 individuals). **c** Ability of young Nile crocodiles to detect a sound stimulus (synthetic buzz) as a function of Signal-to-Noise Ratio (SNR) and loudspeakers spacing (120 trials on 2 individuals).



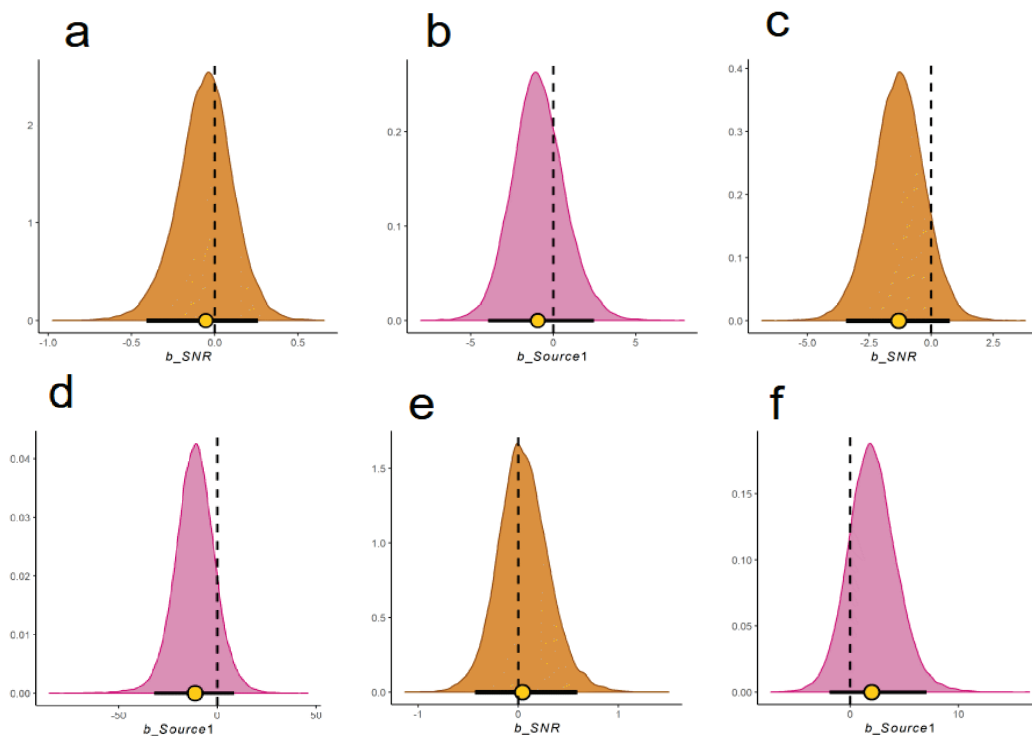
**Supplementary Figure III.2. Bayesian model outcome for the Experiment 1.** Posterior density distributions for the effect of the SNR (a) and target position (b) on the behavioral score probability. Posterior density distributions for the effect of separation angle (c) on the behavioral score probability. Circles show posterior median distribution of the beta-coefficient with 95% credible intervals in black horizontal bars. Vertical dashed line represents a null effect, a positive effect size indicating an increased behavioral score/detection probability.



**Supplementary Figure III.3. Bayesian model outcome for the Experiment 2.** Posterior density distributions for the effect of the SNR (a) and target position (b) on the detection probability. Posterior density distributions for the effect of separation angle (c) on the detection probability. Circles show posterior median distribution of the beta-coefficient with 95% credible intervals in black horizontal bars. Vertical dashed line represents a null effect, a positive effect size indicating an increased behavioral score/detection probability.

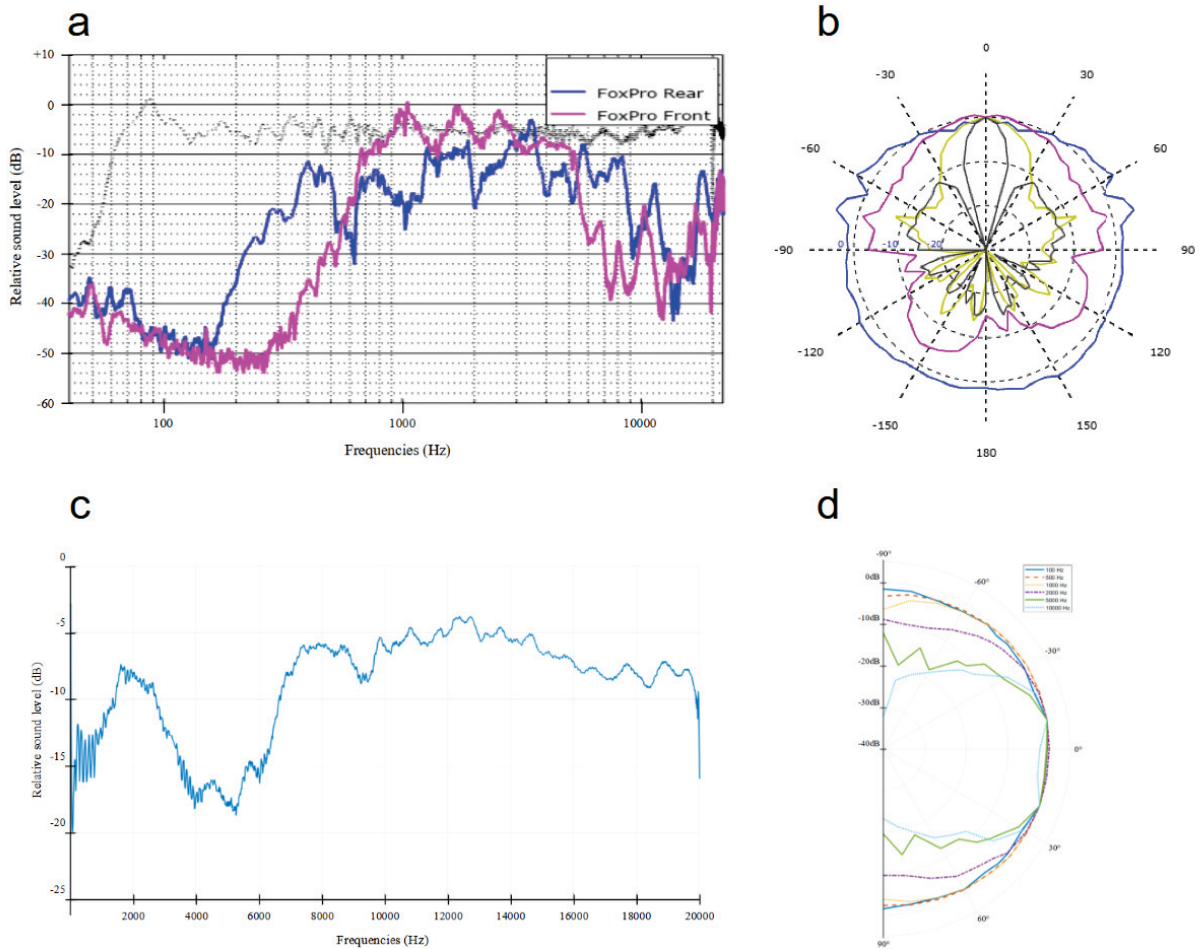


**Supplementary Figure III.4. Bayesian model outcome for the Experiment 3.** Posterior density distributions for the effect of the SNR (a) and target position (b) on the detection probability. Posterior density distributions for the effect of separation angle (c) on the detection probability. Circles show posterior median distribution of the beta-coefficient with 95% credible intervals in black horizontal bars. Vertical dashed line represents a null effect, a positive effect size indicating an increased behavioral score/detection probability.



**Supplementary Figure III.5. Bayesian model outcome: posterior density distributions for the effect of the SNR (a, c, e) and target position (b, d, f) on the reaction time.** The panel a and b refer to the experiment 1, the panel c and d refers to the Experiment 2 and the panel e and f refers to the experiment 3. Circles show posterior median distribution of the beta-coefficient

with 95% credible intervals in black horizontal bars. Vertical dashed line represents a null effect, a positive effect size indicating an increased behavioral score/detection probability.



**Supplementary Figure III.6.** **a** Frequency response of the FoxPro Fusion speakers (rear loudspeaker) used in Experiments 1 and 2. **b** Directivity pattern of the FoxPro Fusion speakers (rear loudspeaker) used in Experiments 1 and 2. **c** Frequency response of the AudioPro, Bravo Allroom Sat speakers used in Experiment 3. **d** Directivity pattern of the AudioPro, Bravo Allroom Sat speakers used in Experiment 3. All measurements were done in a semi-anechoic chamber with calibrated chains.





## 5.2 Supplementary Tables

Tested female → SNR (dB) ↓	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
-20					×											
-18					×	×	×	×	×	×	×	×	×	×	×	×
-16			×	×	×	×	×	×	×	×	×	×	×		×	×
-14	×	×			×	×	×	×		×		×	×		×	×
-12					×		×			×			×			
-10					×		×			×			×			
-8							×			×						
-6							×			×						
-4							×			×						
-2							×									

**Supplementary Table III.1. Details of the trials performed during the experiment 1** (N = 16 female Jacare caimans). A cross "x" indicates that the tested female was challenged by both a co-located and separated stimulus at the corresponding SNR. An empty cell indicates that the female was neither challenged by a co-located nor a separated stimulus.

Tested crocodile → SNR (dB) ↓	1	2	3	4	5	6	7	8
-33					0/1			
-32				0/1	0/1	0/1		
-31								0/1
-30				1/0	0/1			
-29			0/1	0/1				
-28			0/2				1/0	
-27		1/0	1/0			0/1	0/1	
-26			1/0	1/1	0/1			
-25	0/1	1/0		0/1		0/1	0/1	
-24	0/2				1/0			
-23	0/2		1/1		1/0			0/1
-22	1/0				0/1	1/0	1/0	
-21		1/1	0/1	1/0	1/1	1/0	0/1	
-20		0/1			1/1			2/1
-19		1/1				0/1	0/1	
-18	3/0			1/0				0/1
-17				0/1		1/1		
-16		1/0	1/0		1/0			0/1
-15			0/1					0/1
-14								
-13							1/0	

**Supplementary Table III.2. Details of the trials performed during the experiment 2** (N = 8 young Nile crocodiles). Each cell reports the number of stimuli (crocodile calls) played in the co-located and in the separated condition (left and right numbers, respectively).

Tested crocodile → SNR (dB) ↓	1	2
-32	0 / 2	1 / 0
-30	0 / 2	1 / 1
-28	0 / 5	2 / 4
-26	2 / 2	0 / 4
-24	4 / 4	1 / 8
-22	4 / 4	3 / 12
-20	6 / 6	4 / 7
-18	4 / 3	5 / 4
-16	4 / 2	3 / 4
-14	0 / 0	0 / 0
-12	0 / 1	0 / 1

**Supplementary Table III.3.** Summary of all testing conditions for each individual considered in the experiment 3 (i.e. two conditioned juvenile Nile crocodiles). In each cell of the table, the left and right numbers correspond respectively to the number of target signals played in the co-located and in the separated condition.





# IV. Sound categorization

## 1. Ethological experiments

### Sound categorization by crocodilians

This article is in preparation for the *Journal of Experimental Biology*.

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

‡ Co-last authors

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

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

## Abstract

Rapidly detecting and reacting to the information contained in the stream of stimuli to which sensory systems are subjected is a major and necessary challenge for animals, including humans. One cognitive mechanism for achieving this goal is categorization, where the receiving individual considers a continuous variation of a stimulus as belonging to discrete categories. Here we test whether crocodilians are able to partition an acoustic continuum into meaningful categories. Using playback experiments, we first demonstrate that young crocodiles confronted with a continuum of acoustic chimeras ranging from a frog call to a crocodile call classify each acoustic variant into one of these two categories, thus establishing a sharp meaningful boundary where no acoustic boundary exists. With go/no-go experiments, we then observe that this category boundary is likely to be shifted along the acoustic continuum by learning. We further demonstrate that crocodilians rely on the spectral envelope of sound signals to categorize stimuli while they ignore the pitch and the aperiodic component of the signals. This study suggests that sound categorization in crocodilians is a pre-wired faculty that is mobilized in the face of an acoustic continuum, allowing rapid decision making, and highlights the learning-dependent plasticity involved in defining the boundary between sound categories.



# Graphical abstract

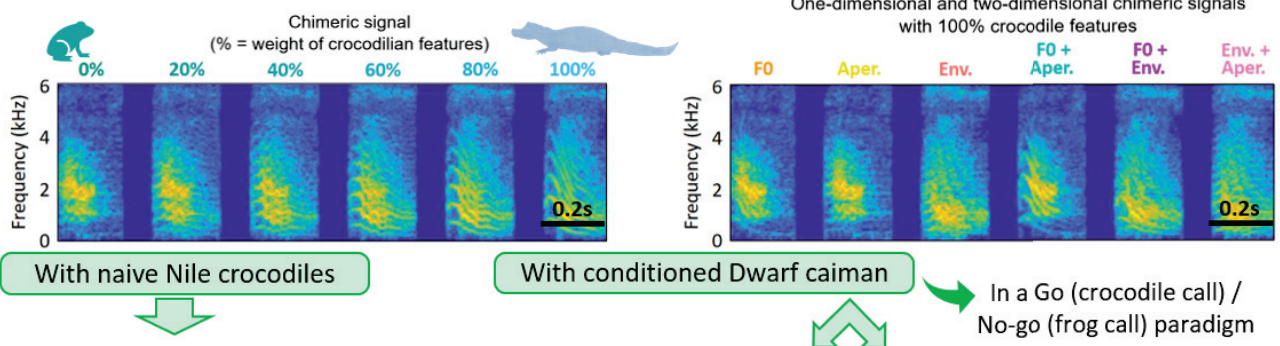
## Sound categorization by crocodilians

Once crocodiles are able to detect and localize sounds, what makes a sound interesting enough for the crocodile to decide to respond to it? Do crocodilians can categorize an auditory continuum?

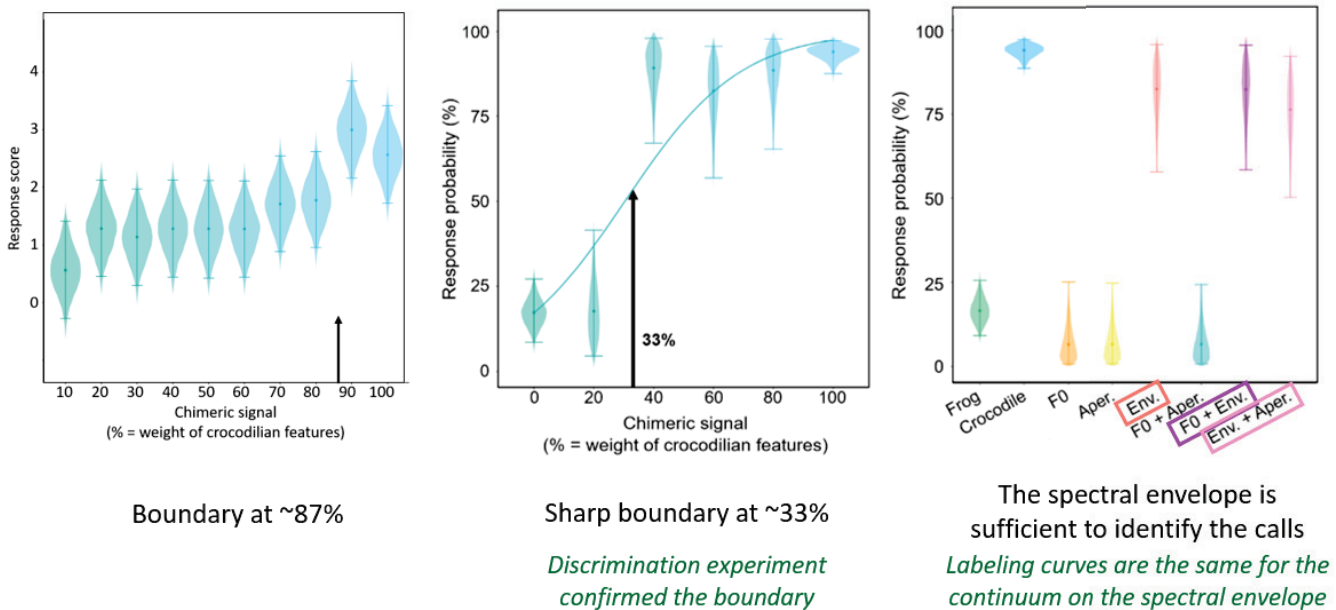
Categorizing sounds, i.e. deciding which stimulus corresponds to which object in the sound scene, such as a conspecific call rather than a call from another species, can be complicated especially when the two signals are acoustically close. The optimal decision in terms of the probability of responding to a conspecific call (like a crocodile contact call) rather than a heterospecific call (a frog call) can be predicted by the acoustic cues carried by the sound stimulus. An acoustic feature threshold value - or informational threshold - can be identified, determining when animals decide on their categorical classification and behavioral response.



### Labelling experiments



### Results



### Conclusion

Young crocodilians faced with an acoustic continuum ranging from a frog call to a crocodile call classify each acoustic variant into one of these two categories on the basis of their spectral envelope, establishing a clear meaningful boundary that is likely to be shifted along the continuum by learning. This study suggests that sound categorization in crocodilians is a pre-wired faculty that can be mobilized for rapid decision making, and highlights the learning-dependent plasticity involved in defining the boundary between sound categories.

## 1.1 Introduction

Detecting and quickly processing information from the massive flow of information emanating from the auditory scene is a major challenge for animals (Hulse, 2002). Although an auditory stimulus can be accurately described physically, its perception by animals depends primarily on their auditory abilities and perceptual processes (Guilford & Dawkins, 1991; Miller & Bee, 2012; Rowe, 2013; Bee & Miller, 2016). Being able to categorize stimuli can help process auditory information and make decisions about how to respond. In this process, animals integrate continuous variation in a sound stimulus as belonging to discrete categories carrying different information. This results in a peak in discrimination between stimuli of different categories versus stimuli of the same category (Wood, 1976; Harnad, 1987).

Categorizing sounds, i.e. deciding which stimulus corresponds to which object in the sound scene, such as a conspecific call rather than a call from another species, can be complicated especially when the two signals are acoustically close. As suggested by the signal detection theory (Bradbury & Vehrencamp, 2011), the optimal decision in terms of the probability of responding to a conspecific call rather than a heterospecific call can be predicted on the one hand by the relative probability of each of the two species being present at the location of the sound source, and on the other hand by the acoustic cues carried by the sound stimulus. An acoustic feature threshold value - or informational threshold- is identified, determining the point at which animals decide on their categorical classification and behavioral response (Weber, 1834; Akre & Johnsen, 2014).

Previous works have shown that frequency parameters are often used to quickly classify sounds into categories. The cricket *Teleogryllus oceanici* categorizes pulsed sounds into attractive versus repulsive when they cross a frequency threshold of 16kHz (Wytenbach et al., 1996). The great tits *Parus major* bases its classification of song notes on their fundamental frequency (Weary, 1990). The Japanese macaque *Macaca fuscata* relies on the temporal position of a frequency peak to classify the contact calls of its species into two functional categories (May et al., 1989). Categorical thresholds may also be related to the temporal characteristics of sound stimuli. Thus, the duration of song notes and calls is the criterion used to establish the threshold between two informational categories and guide the behavioral response of the swamp sparrow *Melospiza georgiana* (Nelson & Marler, 1989) and the mouse *Mus domesticus* (Ehret, 1992) respectively. A recent study showed that Japanese macaques can discriminate conspecific



vocalizations and categorically identify conspecific based on the resonance of the vocal tract characteristics (Furuyama et al., 2017).

Despite these results, our understanding of sound categorization processes in animals remains limited. Among tetrapods, only a few species of birds and mammals have been studied (Fischer, 2006; Green et al., 2020). In the present paper, we present experiments with crocodilians aimed at identifying their ability to form functional categories from continuous sets of stimuli. Crocodilians are phylogenetically related to birds and show common behaviors in this group, such as parental care. While more distant from mammals, crocodilians exhibit many ecological traits that have evolved convergently with that group (Reber, 2020). Despite a number of recent studies, little is known about how crocodilians apprehend their sound world.

In addition to being attentive to their sonic environment to identify potential prey, crocodilians actively use acoustic communication in their social interactions. This channel is essential in their early years where a repertoire of calls is used by the young to communicate with each other and with their parent (A. L. Vergne et al., 2009). Their position as top predators and their auditory abilities suggest that crocodilians are super-efficient and quick to analyze their sound scene, derive relevant information, and make appropriate decisions in terms of behavioral response (Grigg & Kirshner, 2015). For example, crocodiles are able to accurately localize a sound source (Bierman & Carr, 2015; Papet et al., 2019, 2020) and use this ability to effectively perform spatial unmasking to identify a relevant sound despite background noise (Thévenet et al., 2022b). In this work, we test the hypothesis that crocodilians are able to partition an acoustic continuum into categories and identify the acoustic features used for this categorization. Being able to form meaningful categories from auditory scenes should allow crocodilians to facilitate rapid decision making, and adjust a behavioral response such as an approach to the sound source accordingly.

In a first experiment, we test whether this ability to partition an acoustic continuum into discontinuous categories is innate by observing the response of naive crocodiles to frog and crocodile calls as well as to chimeric signals containing a variable proportion of frog and crocodile acoustic features. Second, we test whether the boundary between these categories can be modulated by learning by training caimans to approach a speaker emitting a crocodile call and ignore the frog call (go/no-go procedure) and then testing them with a continuum of acoustic chimeras. Finally, we investigate the acoustic basis of this categorization using acoustic chimeras based on only certain dimensions of the signal (pitch, aperiodic component,

spectral envelope). Our results establish that crocodylians follow behavioral decision rules based on the establishment of sound categories constructed from the spectral envelope of signals.

## 1.2 Methods

### 1.2.1 Experiment 1: Labeling experiment with untrained naïve crocodylians

#### *Animals and experimental setup*

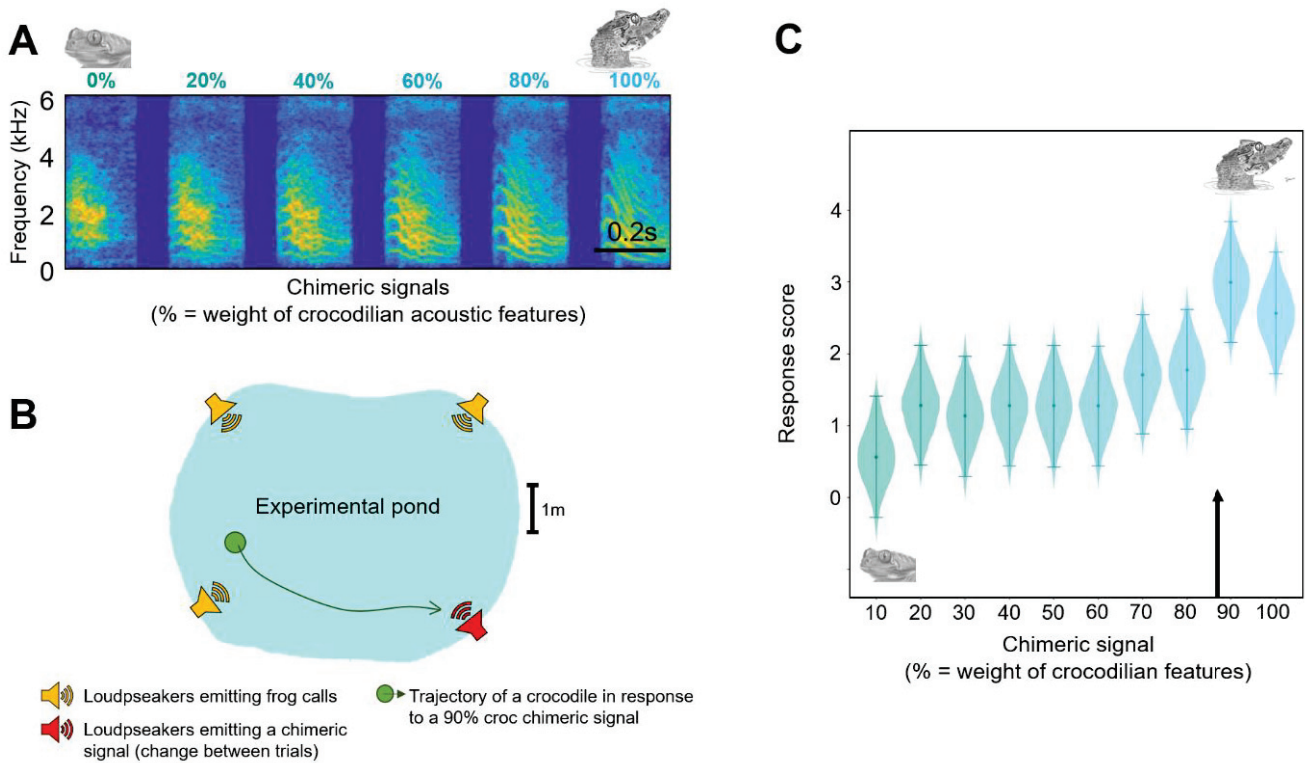
The experiment was performed at the "Crocoparc" zoo (Agadir, Morocco). We worked with captive-born juvenile Nile crocodiles *Crocodylus niloticus* (n = 14 individuals; sex unknown; aged between 1 and 2 months). These animals were housed together in an outdoor enclosure, not visible to the public. They had never been included in any experiment before. We conducted the experiments outdoors, at night, in an artificial tank of approximately 40 m<sup>2</sup> (maximum dimensions: 6 X 7 meters, Figure IV.1B). Four loudspeakers were placed near the water on the edge of the basin at four opposite locations (FoxPro Fusion© with Visaton SL 87 ND internal speakers; see appendix Supplementary Figure IV.1 for technical specifications).

#### *Acoustic stimuli*

We tested the crocodiles with an acoustic continuum between two signals that, while sharing some acoustic similarities, came from two different animal species and thus had different biological meanings: a Saharan frog call *Pelophylax saharicus*, a species naturally present in the zoo enclosure and routinely participating in the night time sound scene, and a contact call of a young Nile crocodile, which previous studies have shown to be attractive to crocodiles (A. L. Vergne et al., 2012). This frog call and this crocodile call are harmonic series in the same frequency range ([1000, 3500] Hz and [250, 4000] Hz respectively) and are of comparable duration (190 ms ± 17ms ; Figure IV.1A). To limit pseudo-replication, we used 4 frog calls (recorded by Léo Papet in the zoo) and 4 juvenile crocodile calls (recorded during a field expedition in the Okavango Delta by N. Mathevon and T. Aubin). Each of the 4 acoustic continua used during the experiment was made from one of these frog calls and one of these crocodile calls.







**Figure IV.1. Categorization tests with naive crocodilians reveal that they switch their behavioral response across a specific acoustic boundary along a signal continuum (Experiment 1).** **A** Acoustic continuum of frog/crocodile chimeric signals, ranging from a ‘frog’ signal (100% frog / 0% croc) to a ‘crocodile’ signal (0% frog / 100% croc). The ‘frog’ signal was used to create a frog chorus. Only 5 of the 10 signals tested are shown here (the entire continuum spanned from a 100% frog / 0% crocodile signal to a 0% frog / 100% crocodile signal, with 10% increments; see Methods for details of the acoustic morphing procedure). **B** Experimental set-up. Throughout each experiment (average duration = 2.5 hours), all four speakers were emitting 100% frog calls (1 call every 9 to 13 seconds, different rhythm for each speaker, reproducing a frog chorus). During each test, one of the speakers switched to one of the chimeric signals (duration of a test = 4 minutes; tests separated by at least 10 minutes; different test speaker from one test to another; see Methods for details). The crocodile’s response to the chimeric signal was assessed for 10 minutes after the start of each test. Each individual (N = 14 Nile crocodiles, 1-2 months old) was tested successively with all chimeric signals, presented in random order during the experiment. **C** Behavioral response of crocodiles to chimeric signals. Solid circles show the fitted probabilities of getting a response (the crocodile approached the test speaker) expressed as the median of posterior distribution with 95% CI. Violin plots show the distribution of the fitted values for each signal. The abrupt increase in approaches to the test speaker from the 10% frog / 90% croc signal (marked by the orange arrow) suggests sound categorization. This result was obtained with naïve, untrained crocodiles, which had therefore never heard the test signals before the experiment.

To create the acoustic continua (acoustic morphing), we first decomposed the frog and crocodile calls into three orthogonal acoustic dimensions: the pitch (fundamental frequency)  $F_0$ , the aperiodicity  $A_{per}$  and the spectral envelope  $Env$ . When possible, this decomposition was performed with the open source Matlab toolbox STRAIGHT (Kawahara et al., 1999; Kawahara, 2006) which is generally dedicated to analyze and synthesize human speech sounds. For some limitation in the parameters' variation, STRAIGHT failed to extract the pitch contour of the frog calls. To do so, we used an auto-correlation to compute the modulation frequency of the temporal envelope of the carrier frequency in each temporal window. This homemade code is provided as an open source resource. Each acoustic chimera was then re-synthesized using STRAIGHT by combining the three dimensions of a frog call and those of a crocodile call in different proportions. The aperiodic and the spectral envelope dimensions from frog and crocodile calls were linearly combined while their pitch were combined on a logarithmic scale. Finally, the continuum consisted of 11 calls, from a 100% frog call to a 100% crocodile call, through 9 chimeric calls where the relative proportion of the three dimensions between the two species varied from 90% frog / 10% crocodile to 10% frog / 90% crocodile, with a 10% step between each signal (Figure IV.1A).

### ***Experimental procedure***

The experiments took place at night, starting one hour after sunset. It is during the night that the frogs sing the most and the crocodiles are the most active. To avoid habituation, each individual was tested in only one experimental session (one night), with a single set of acoustic chimeras. The tested crocodile was released into the basin three hours before the start of the experiment. From sunset until the end of the experiment, the 4 speakers emitted frog calls, in an unsynchronized manner (i.e. every 9, 10, 11 and 13 seconds for speakers  $n^{\circ}1$  to  $n^{\circ}4$  respectively). This setup allowed to mimic a frog chorus, to which crocodiles are used to. Then, the crocodile was exposed to a succession of experimental tests using each acoustic chimera of the continuum (see Figure IV.1B for a schematic of the experimental setup). All sound signals used in the experiment were emitted at an intensity of 67 dBA at 1 m from the speaker.

At the time of a test, a loudspeaker was randomly selected, eliminating the one or ones from which the crocodile was closest. For 2 minutes, this test speaker continued to emit frog calls, but with a rhythm of 1 every 12 seconds. The choice of a fixed rate of emission, the same for all tests of all crocodiles, eliminated the risk of obtaining different behavioral responses following different rates of emission. At the end of the first 2 minutes, the test speaker started to emit for the next 2 minutes one of the acoustic chimeras (chosen between the 90% frog / 10%



crocodile and the 0% frog / 100% crocodile) at the same rate of one call every 12 seconds. At the end of the 2 minutes, the test speaker resumed emitting frog calls at its original rate. The next test was initiated at least 10 minutes later, from a different speaker. The behavior of the crocodile was observed and evaluated independently by 4 observers for 10 minutes from the start of the test. The entire experiment was filmed by an infrared camera (ABUS TVCC34010).

### ***Analysis of behavioral responses***

The intensity of the crocodiles' behavioral response to sound stimuli was scored on a scale of 0 to 5 by the observers (0: no response; 1: set in motion without swimming; 2: swimming without preferred direction; 3: swimming in the direction of the speaker but not exceeding 1/3 of the distance to the speaker; 4: swimming in the direction of the speaker without approaching the speaker within 50 centimeters; 5: swimming in the direction of the speaker and approaching within 50 centimeters). For each test, the median of the scores given by the observers was calculated.

Statistical analyses were performed in R (v3.6.2) using Bayesian mixed models fitted with the brms R package (Bürkner, 2017). The advantages of using the Bayesian approach are multiple, including its high flexibility, quantification of uncertainty in estimates, and intuitive interpretation of confidence intervals (J. Kruschke, 2014; J. K. Kruschke & Liddell, 2018). Response scores were modeled using a gaussian function. The model included subject-specific random intercepts. 5000 iterations were run over four MCMC chains with the first 500 iterations of each chain used to adjust the algorithm. Because the behavior of the tested animals was difficult to predict, flat priors were kept for this model. Results were summarized as medians of the posterior distributions and 95% credible intervals (CIs). Credible intervals for estimates that do not include zero indicate a credible effect given the observed data and model structure (J. K. Kruschke & Liddell, 2018). Similarly, when contrasting two conditions, CIs excluding the null value can be inferred to indicate a credible difference between the conditions.

## **1.2.2 Experiment 2: Labeling experiments with trained crocodilians**

### ***Animals and experimental setup***

The experiment was conducted at the ENES laboratory. We worked with young dwarf Cuvier's caiman *Paleosuchus palpebrosus*, born in captivity at the "Zoo de Paris" (N = 5, including 2 females and 3 males; age: 2.5 years; size: 59 ± 2 cm). These animals had never been included

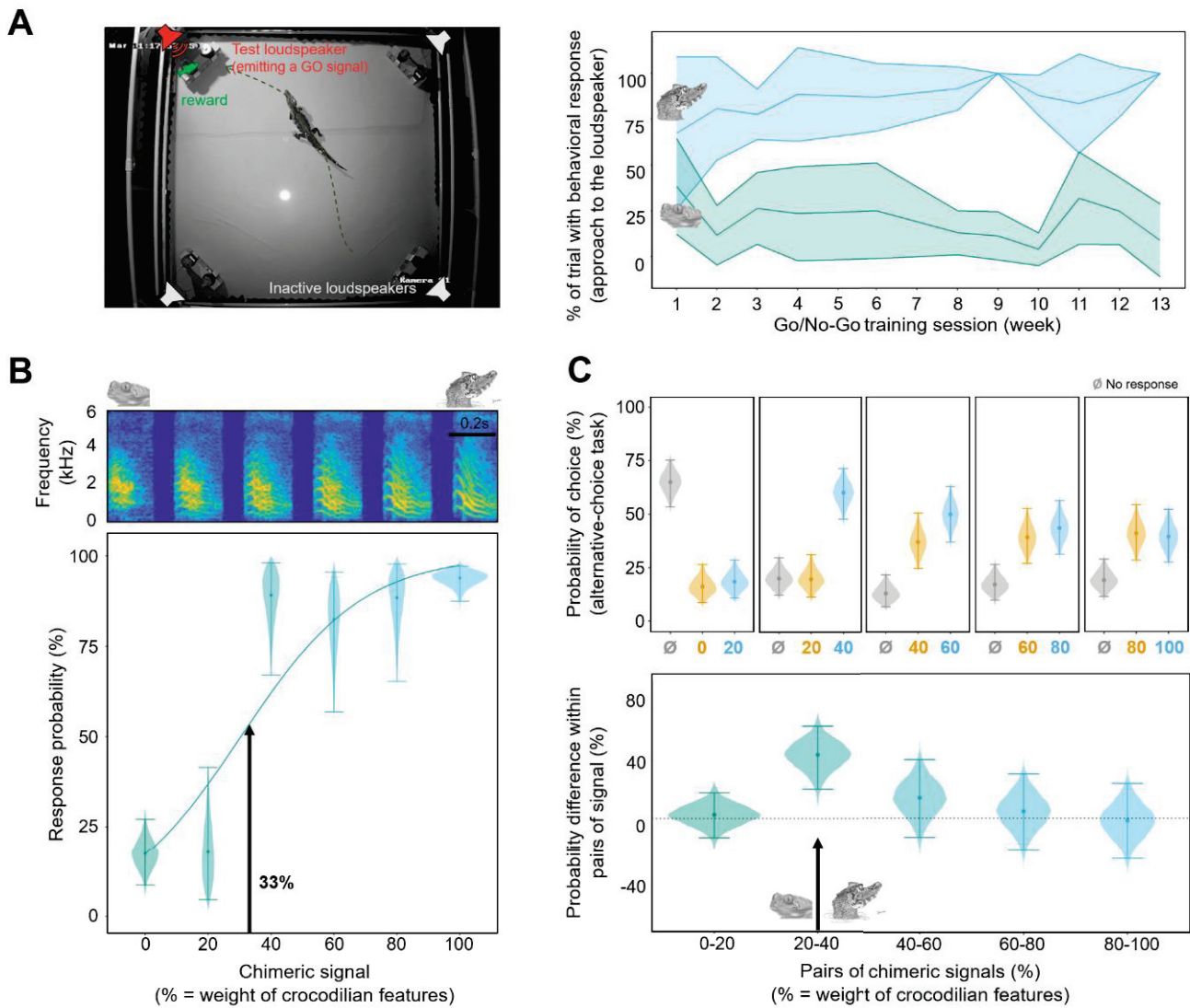
in an experimental protocol prior to the experiment. Individuals were housed all together in a large enclosure (16 m<sup>2</sup>). The room was maintained at  $28 \pm 1$  °C, with a light on from 9:00 am to 9:00 pm. The experiment was conducted in a square-shaped pool (1.75 m wide), placed in a soundproof cabin, with a water depth of 12 cm allowing the animals to swim freely (water temperature: 29°C). Four speakers (Audiopro Bravo Allroom Sat, see appendix Supplementary Figure IV.1 for technical specifications) were placed just above the water surface at the four corners of the pool (Figure IV.2A). Sound playback was controlled by the experimenter from outside the booth using a computer delivering the signals to an amplifier (Yamaha AX-397) connected to the speakers. A remote-controlled reward system was associated with each speaker, which allowed for multiple delivery of food rewards during a single test session. All tests were performed in low light conditions to limit the stress of the animals and increase their motivation to respond (these animals are more active at night). The whole experiment was filmed (ABUS TVCC34010 camera), which allowed real time and delayed observations.

### *Acoustic stimuli and experimental procedure*

The ability of animals to categorize sound signals forming an acoustic continuum between a “frog call” and a “crocodile call” was tested here via a go/no-go experiment. The stimuli used were identical to those used in Experiment 1 except that we reduced the step between each signal: the continuum consisted of 6 calls, from a 100% frog call to a 100% crocodile call, through 4 chimeric calls where the relative proportion of the three dimensions between the two species varied from 80% frog / 20% crocodile to 20% frog / 80% crocodile, with a 20% step between each signal (Figure IV.2B). They were emitted at a sound level of 60 dB SPL at 1m from the speaker.

In a first step, individuals were trained to respond positively (i.e. by approaching the speaker) to a crocodile call. Training took place in weekly sessions via operant conditioning using a food reward. The subject was first placed in the pool and did not receive any stimulation for an acclimatization period of at least 3 min (we always waited until the animal had stopped exploring the pool). Then, the GO stimulus (sequence of crocodile calls) and the NOGO stimulus (sequence of frog calls) were presented three to seven times each in a random order (10 calls per sequence at a rate of one call every  $2.5 \pm 0.4$  sec, total sequence duration = 35 sec;  $10 \pm 3$  sequences in total per training session). The animal was systematically rewarded if it came within 35 cm of the active speaker (corresponding to the location of the food reward system) when a sequence of crocodile calls (GO stimulus) was presented. There was a minimum delay of 1 min 30 between two trials.





**Figure IV.2. Crocodilians trained to respond to a meaningful signal highlight their ability to form a sensory boundary along an acoustic continuum.** **A** Training procedure: Young caimans (N = 5) were trained using food reward to approach the speaker only when it emits a 100% croc signal (GO-NOGO procedure). As the graph on the right shows, training reinforces the natural propensity of animals to respond to that signal. **B** Experiment 2: Categorization tests along the acoustic continuum suggest a sensory boundary between the 20% and 40% croc signals (black arrow). Solid circles show the fitted probabilities of getting a response (the crocodile approached the test speaker) expressed as the median of posterior distribution with 95% CI. Violin plots show the distribution of the fitted values for each signal. **C** Experiment 3: Fitted probability of choosing between two chimeric signals in an alternative-choice task (upper graph), and difference between the probabilities measured for each pair of signals (below). Choice between neighboring chimeric signals is effective only between the 20 and 40% croc signals (black arrow).

Once trained, i.e., when the subject responded correctly in more than 75% of the tests, the subject moved on to weekly test sessions, during which the animal was confronted with three to five sequences of calls, randomly selected from the acoustic continuum. No rewards were given during the tests when the crocodile response was measured, regardless of the signal and the animal's response. However, in order to maintain motivation over the test sessions, the animal's response was regularly reinforced by repeating the go/no-go protocol, i.e. by associating a positive response to the GO signal with a food reward. These reinforcements were randomly distributed during the sessions. During each test session, the caiman heard  $12 \pm 3$  sequences of calls, including 40% test stimuli (without reward), 40% GO stimuli, 20% NOGO stimuli. Testing sessions were repeated until each caiman had been tested three times with each of the test stimuli of the continuum (i.e. total of  $12 \pm 4$  test stimuli per caiman throughout the experiments, with at least 3 repetitions for each of the 4 test stimuli).

### *Analysis of behavioral response*

The response of the animals to the acoustic stimuli was measured from the videos using Kinovea software, v0.8.24 Beta (Puig-Diví et al., 2019), during the 45 seconds following the emission of the first call of the tested sequence. For each trial, the response was scored as "0" (no approach) or "1" (reach the food reward system within 35 centimeters of the speaker). We also measured the latency between the first call in the sequence and the onset of the behavioral response (first movement), if any.

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

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

### *Experimental procedure*

After completing Experiment 2, the same caimans were tested in another experimental paradigm, also based on the GO-NOGO approach, but in a two-choice testing procedure. The



goal was to test the ability of crocodilians to categorize acoustic signals when confronted with two sound sources alternately emitting a pair of neighboring acoustic chimeras within the continuum. 5 pairs of stimuli were formed: 100% frog / 0% crocodile with 80% frog / 20% crocodile, 80% frog / 20% crocodile with 60% frog / 40% crocodile, 60% frog / 40% crocodile with 40% frog / 60% crocodile, 40% frog / 60% crocodile with 20% frog / 80% crocodile, and 20% frog / 80% crocodile with 0% frog / 100% crocodile. The two stimuli of a pair were presented antiphonally during 35 seconds through two speakers located at equal distance from the animal. The overall rhythm of emission was the same as in the second experiment, which means that only 5 calls were emitted per speaker instead of 10 calls.

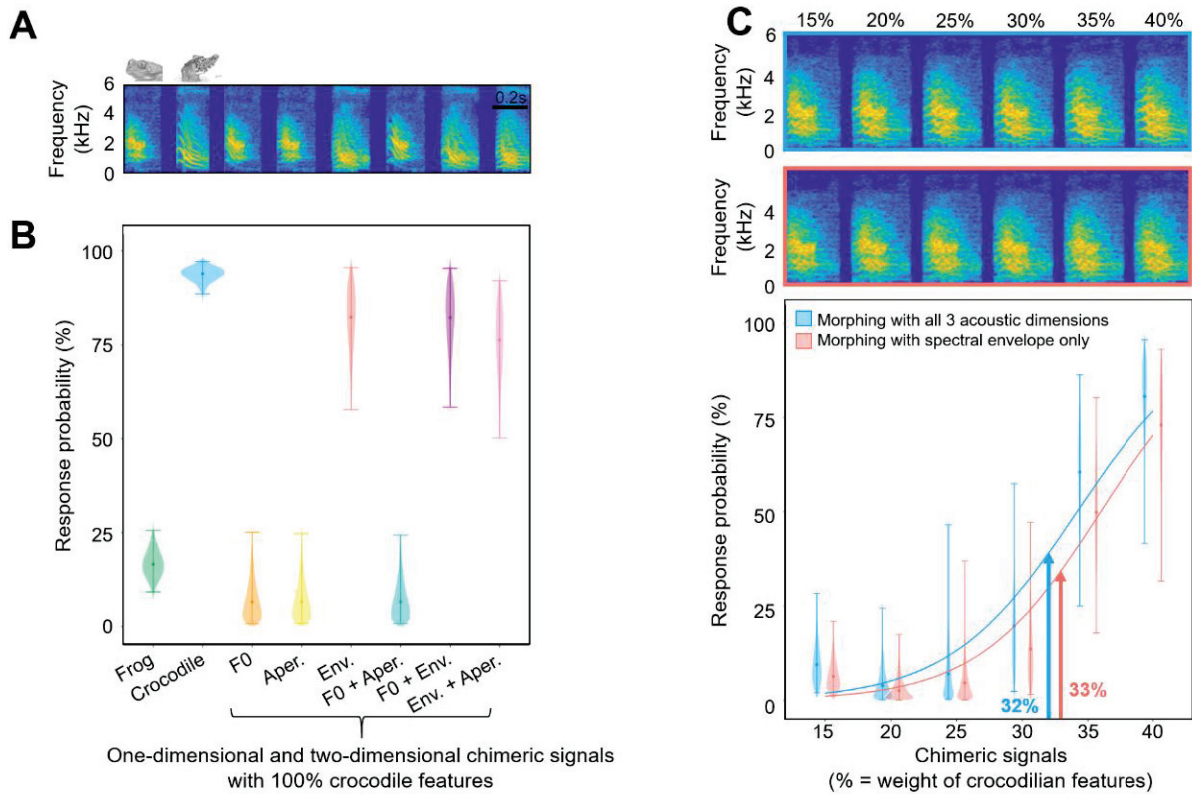
A test session always began with a training pair combining a 100% frog / 0% crocodile signal (NOGO) with a 0% frog / 100% crocodile signal (GO). Discrimination tests only began if the animal moved to the speaker emitting the GO signal twice in a row. If the animal made the wrong choice, it was trained again with the GO stimulus. Then, the animals were tested with all the five pairs of stimuli beginning once by one stimuli of the pair and then the other, in a random order (i.e. 10 pairs of test stimuli were presented by session). In total, each caiman was tested with  $87 \pm 31$  pairs of stimuli throughout the experiment (or  $87 \pm 3$  times each pairs of stimuli, all caimans combined).

### *Analysis of behavioral response*

For each trial, the caiman was considered to have made its choice between the two speakers when it approached one of them within 35 cm within 45 seconds of the call. Choices were modeled using a multinomial function (categorical family), with random intercepts per animal.

## **1.2.1 Experiment 4: Identification of acoustic cues supporting category labeling**

This experiment involved the same caimans, after they had performed experiments 2 and 3. To investigate which signal dimensions (F0, A<sub>per</sub>, and/or Env) are used by crocodilians to categorize calls as "frog" or "crocodile", we first tested the caimans with acoustic chimeras bearing on only one or two dimensions. In this first step, each chimera was constructed based on a 100% frog / 0% crocodile call, and one or two dimensions were raised to its value in the 100% crocodile call. In addition to the two control calls 100% frog / 0% crocodile and 0% frog / 100% crocodile, the following 6 acoustic chimeras were tested (Figure IV.3A):



**Figure IV.3.** The categorization of frog/croc chimeric signals by crocodilians is based on the spectral envelope of the sound stimulus, excluding other acoustic dimensions (Experiment 4). **A** One-dimensional and two-dimensional chimeric stimuli were created from a 100% frog call by transforming into 100% croc three acoustic dimensions separately or in pairs: pitch (F0), aperiodic dimension (Aper), and spectral envelope (Env), and their combinations F0 + Aper, F0 + Env, and Env + Aper. **B** Results of playback experiments with animals trained to respond to crocodile call. Solid circles show the fitted probabilities of getting a response (the crocodile approached the test speaker) expressed as the median of posterior distribution with 95% CI. Violin plots show the distribution of the fitted values for each signal. Only stimuli with 100% crocodile spectral envelope elicited a behavioral response. The subjects did not respond to stimuli with only 100% F0 croc or 100% Aper croc. **C** When the spectral envelope is the only dimension to change in the acoustic continuum (in pink), the sensory boundary remains close to that identified for the original continuum where all three dimensions of the signal change together (boundaries between the 32 and 33% croc signals: blue and pink arrows).





- 0% F0, 100% Aper, 100% Env frog / 100% F0, 0% Aper, 0% Env crocodile;
- 100% F0, 0% Aper, 100% Env frog / 0% F0, 100% Aper, 0% Env crocodile;
- 100% F0, 100% Aper, 0% Env frog / 0% F0, 0% Aper, 100% Env crocodile;
- 0% F0, 0% Aper, 100% Env frog / 100% F0, 100% Aper, 0% Env crocodile;
- 0% F0, 100% Aper, 0% Env frog / 100% F0, 0% Aper, 100% Env crocodile;
- 100% F0, 0% Aper, 0% Env frog / 100% F0, 100% Aper, 100% Env crocodile;

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

Since the results of this first step showed that the spectral envelope was a decisive acoustic criterion for categorization (see Results below), we tested whether the boundary between the two categories "frog" and "crocodile" was identical between a control continuum where the 3 acoustic dimensions varied and a continuum where only the spectral envelope varied. Based on the results of experiments 2 and 3 which indicated that the acoustic boundary should be for a signal containing around 30% crocodile features (see Results below), we created a series of 6 acoustic chimeras where the 3 dimensions of the signal varied together, from a 85% frog / 15% crocodile signal to a 60% frog / 40% crocodile signal, with a 5% step (Figure IV.3C). We also created another series of 6 acoustic chimeras, where only the spectral envelope was involved, from a 100% F0, 100% Aper, 85% Env frog / 0% F0, 0% Aper, 15% Env crocodile signal to a 100% F0, 100% Aper, 60% Env frog / 0% F0, 0% Aper, 40% Env crocodile signal. The experimental and behavioral response analysis procedures were identical to those in Experiment 1. In total, the caimans were tested with  $54 \pm 2$  test stimuli throughout the experiments.

## 1.3 Results

### 1.3.1 Experiment 1: Naïve crocodilians respond discontinuously to continuous stimuli

As illustrated in Figure IV.1C, crocodiles confronted with chimeric frog/crocodile calls are attracted to the speaker when the stimulus has a very high proportion of crocodile features. Specifically, their response score increases sharply for the 10% frog / 90% crocodile stimulus

whereas it is low for stimuli with a lower proportion of crocodile features (response score 1.2 point higher for the 90% crocodile signal compared to the 80% crocodile signal, 95% CI [0.1, 2.3], 98.5% of the posterior distribution is positive).

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

As shown in Figure IV.2A, the caimans learned to systematically move toward the speaker emitting the 100% crocodile signal. At the beginning of training, the response rate is around 67.4%, while it is only 38.3% for the 100% frog signal. After  $2 \pm 1$  weeks of go/no-go training (individual learning curves are presented in appendix Supplementary Figure IV.2), this response rate settles above 75% and the caimans were 76.2% more likely to respond to the GO stimulus versus the NOGO stimulus (95% CI [66.5, 84.1]).

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

The response of the caimans thus appears clearly dichotomous, with a boundary between two signal categories located between the 80% frog / 20% crocodile signal and the 60% frog / 40% crocodile signal (probability of response 69.2% higher for the second category, 95% CI [55.0, 79.5] ; Figure IV.2B). This boundary was remarkably stable across individuals: all animals tested reached a maximum response rate of 33% for the 80% frog / 20% crocodile signal, whereas they all reached a response rate of at least 75% for the 60% frog / 40% crocodile signal. These results were confirmed when comparing the latency to respond to the stimulus between the frog category (100% and 80% frog stimuli, mean latency of  $12.8 \pm 12.3$  s) and the crocodile category (40% crocodile and above stimuli, mean latency of  $5.3 \pm 6.4$  s): the latency decreased by 8.9 s (95% CI [-25.5, -1.3]), while there was no difference of latency for stimuli within the



frog category (difference of -2.0 s, 95% CI [-14.2, 23.9] between the 100% and 80% frog stimuli) and for stimuli within the crocodile category (difference of -1.6 s between the 40% and 60% crocodile stimuli, 95% CI [-16.1, 12.0]; -2.9 s between the 40% and 80% crocodile stimuli, 95% CI [-17.9, 6.8]; and 0.5 s between the 40% and 100% crocodile stimuli, 95% CI [-14.7, 5.6]).

The results obtained with Experiment 3 (alternative choice test, Figure IV.2C) strengthen the conclusions drawn from Experiment 2. Indeed, when presented with a choice between the 80% frog / 20% crocodile signal and the 60% frog / 40% crocodile signal, the caimans show a clear preference for the second signal: they were 40.5% more likely to choose the 40% crocodile signal (95% CI [18.5, 58.7]). Conversely, when confronted with the 100% frog / 0% crocodile and 80% frog / 20% crocodile signal pair, the caimans responded only weakly and equally to both signals (probability of 65.0 to get no response, 95% CI [53.3, 75.2], difference of only 2.3%, 95% CI [-12.3, 16.4] to choose one or another stimulus). For all pairs involving signals containing at least 40% crocodile acoustic features, the caimans showed no preference, responding strongly to either signal in the pair (pair [40%, 60%]: 87.2% of response, 95% CI [78.5, 93.3], probability of a different response between the two signals = 13.0%, 95% CI [-12.4, 37.2]; pair [60%, 80%]: 83.0% (95% CI [73.7, 90.2], probability of a different response = 4.3%, 95% CI [-20.2, 28.1]; pair [80%, 100%]: 80.9%, 95% CI [71.2, 88.6], probability of a different response = -1.4%, 95% CI [-25.3, 22.2]. The latencies to respond were however not different from one pair of stimuli to another (minor difference of 2.8 s, 95% CI [-1.4, 8.8] between the between category pair [20%, 40%] and the frog category pair [0%, 20%]; -1.5 s, 95% CI [-4.4, 0.6] of difference between the pair [20%, 40%] and the crocodile category pairs).

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

Only the acoustic chimeras containing the spectral envelope of the crocodile call induced a significant behavioral response from the caimans (Table IV.1, Figure IV.3B). The presence of the crocodile spectral envelope alone is sufficient to induce a response identical to that obtained with the 0% frog / 100% crocodile signal (negligible difference of response probability of -8.3%, 95% CI [32, 3.9]). Concordant results are obtained if we compare the latency times (Table IV.2).

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% F0 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% Aper crocodile</b>	5.5	-9.6 [-20.5, 8.0]	-87.8 [-94.7, -70.0]
<b>100% F0 + Env crocodile</b>	85.4	69.0 [45.4, 83.1]	-8.3 [-31.5, 4.0]
<b>100% F0 + Aper crocodile</b>	5.5	-9.6 [-20.4, 8.5]	-87.8 [-94.7, -69.5]
<b>100% Env + Aper</b>	78.7	62.4 [37.3, 79.1]	-15.0 [-40.3, 0.7]

**Table IV.1. Effect of the stimulus type on the caimans' behavioral reaction.** Medians of posterior distribution (%) and 95% CI. Evidence for a difference between stimulus can be directly inferred when the 95% CI excludes 0.

Stimuli	Fitted latency (sec)	Median [95% CI] difference with the NOGO stimulus	Median [95% CI] difference with the GO stimulus (sec)
<b>100% frog / 0% crocodile</b>	17.0		<u>10.3 [4.5, 21.4]</u>
<b>0% frog / 100% crocodile</b>	6.7	-10.3 [-21.4, -4.5]	
<b>100% F0 crocodile</b>	20.7	3.0 [-11.1, 35.8]	13.3 [1.5, 49.4]
<b>100% Env crocodile</b>	6.5	-10.2 [-21.9, -2.6]	-0.1 [-4.3, 7.0]
<b>100% Aper crocodile</b>	14.0	-2.8 [-14.6, 14.0]	7.2 [-0.4, 27.1]
<b>100% F0 + Env crocodile</b>	6.8	-9.9 [-21.8, -2.1]	0.1 [-4.1, 7.4]
<b>100% F0 + Aper crocodile</b>	7.9	-8.7 [-20.5, 0.7]	1.2 [-3.5, 11.5]
<b>100% Env + Aper</b>	12.5	-4.4 [-15.5, 8.7]	5.8 [-0.5, 21]

**Table IV.2. Effect of the stimulus type on the caimans' latency time to react.** Medians of posterior distribution and 95% CI. Evidence for a difference between stimulus can be directly inferred when the 95% CI excludes 0.

Comparison of labeling between the control acoustic continuum and the acoustic continuum involving only the spectral envelope shows that caimans establish the boundary between the



two categories "frog" and "crocodile" at the same location on both continua (67.7% frog / 32.3% crocodile and 67% frog / 33% crocodile, respectively). Furthermore, there is no significant difference in behavioral response between full chimeras and chimeras involving only the spectral envelope throughout the acoustic continua (median difference of 5.1 %, 95% CI [-4.6, 15.8], Table IV.3). Consistent results are obtained if latency times are compared (Table IV.3).

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

**Table IV.3. Effect of stimulus type on the caimans' behavioral reaction and latency time.** Medians of posterior distribution (%) and 95% CI. Evidence for a difference between stimulus can be directly inferred when the 95% CI excludes 0.

## 1.4 Discussion

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

Confronted with chimeric signals consisting of a mixture of "frog" and "crocodile" characteristics, the tested individuals made the choice to approach or not the speaker. One

would have expected variations in behavioural response to be linearly correlated with the percentage of "crocodile" signal in the chimeras. This was not the case. While our animals had previously heard frog calls, which are naturally present in the zoo ponds, as well as calls from conspecifics, they had never experienced acoustic chimeras. Our results thus suggest that naive crocodilians practice sound categorization innately.

It is likely that crocodilians can form categories under other circumstances. Because crocodilians have developed a vocal repertoire, it is possible that this ability to categorize is also useful during acoustic communication between individuals. For example, young crocodilians emit cries with "graded" acoustic characteristics by playing on intensity, energy distribution in the frequency spectrum and frequency modulation. A study conducted with the black caiman *Melanosuchus niger* showed that the repertoire of young caimans extends along an acoustic continuum from contact calls allowing the cohesion of the group of young to distress calls inducing maternal protection (A. L. Vergne et al., 2011). It is possible -although not yet demonstrated- that young caiman and their mothers establish two distinct categories along the contact-distress acoustic continuum, which would facilitate decision making. However, it remains to be established where the boundary between contact and distress lies. This boundary is likely to vary according to the context. If the level of risk experienced by the young (in terms of predation pressure for example) is high, it is likely that the boundary between the contact and distress categories may be at a different level than if the level of risk is lower. This hypothesis remains to be tested.

An interesting result of the labelling experiment (Experiment 2) is the evidence of the role of learning in defining the boundary between two sound categories. It may seem surprising that crocodilians trained to respond to a 0% frog-100% crocodile signal have a low categorization threshold (when the signal is "33% crocodile"). This result contrasts with that obtained in the first experiment, where the crocodiles received no prior training - except their own experience in the zoo environment - and were never rewarded. In this first experiment, the young crocodiles only came to the speaker if the signal contained a high proportion of "crocodile" information (around "85% crocodile"). One might have expected that trained individuals would refine their response and that their response threshold would become even higher, since only the "100% crocodile" signal led to a reward. However, it should be kept in mind that in our laboratory conditioning protocol, animals were not confronted with any aversive reinforcer, meaning that there is no risky choice (Bradbury & Vehrencamp, 2011). Choosing to come to the speaker could only lead to two outcomes: being rewarded or getting nothing. Thus, the "punishment"



was light, especially given the short travel distance. This may have led individuals to never take the risk of missing a potentially rewarded signal, and thus be more likely to respond to signals with only a small proportion of crocodile features. In animals, context-dependent decisions are indeed strongly related to estimates of the probabilities of making a good choice (DeGroot, 1970; Akre & Johnsen, 2016).

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

What are the mechanisms underlying this sound categorization process? In particular, can it be considered as categorical perception? Categorical perception is defined as the process by which stimuli are considered similar or different depending on whether or not they fall into the same perceptual category (Harnad, 1987; Green et al., 2020). Strictly speaking, two features characterize a categorical perception: 1) the labelling of stimuli in one or the other of the categories considered, which corresponds to what we observed in our experiments with crocodilians; 2) a "category boundary effect" in which two stimuli belonging to the same category are differentiated less rapidly than two stimuli located on either side of the category boundary. We could not rigorously test this second condition with our crocodiles. Testing the category boundary effect would require numerous Go/No-go tests (aiming to discriminate the two stimuli in each pair of stimuli) that are not reasonably feasible with these animals due to the small number of individuals available in the lab combined with a long learning time (unlike what is possible with humans, birds or monkeys; Morse and Snowdon, 1975; Sinnott et al., 1976; Kuhl and Miller, 1978; Kluender et al., 1987; May et al., 1989; Weary, 1990; Wetzel et al., 1998; Furuyama et al., 2017; Burgering et al., 2019). In our protocol, it was also difficult to use a habituation/deshabituation approach (Nelson & Marler, 1989; Wytenbach et al., 1996; Fischer, 1998) because crocodilians become satiated with food rewards rather quickly and quickly lose motivation to respond. Despite these experimental limitations, our observations strongly suggest the presence of a categorical perception of sound stimuli in the crocodilians tested. Indeed, while caimans were conditioned to respond to the 0% frog - 100% crocodile stimulus in the choice experiment (Experiment 3), they only showed a preference for the 100% crocodile stimulus when the pair "80% crocodile/100% crocodile" was emitted. In other words, the crocodilians tested reacted as if they were truly unable to perceptually differentiate between these two stimuli.

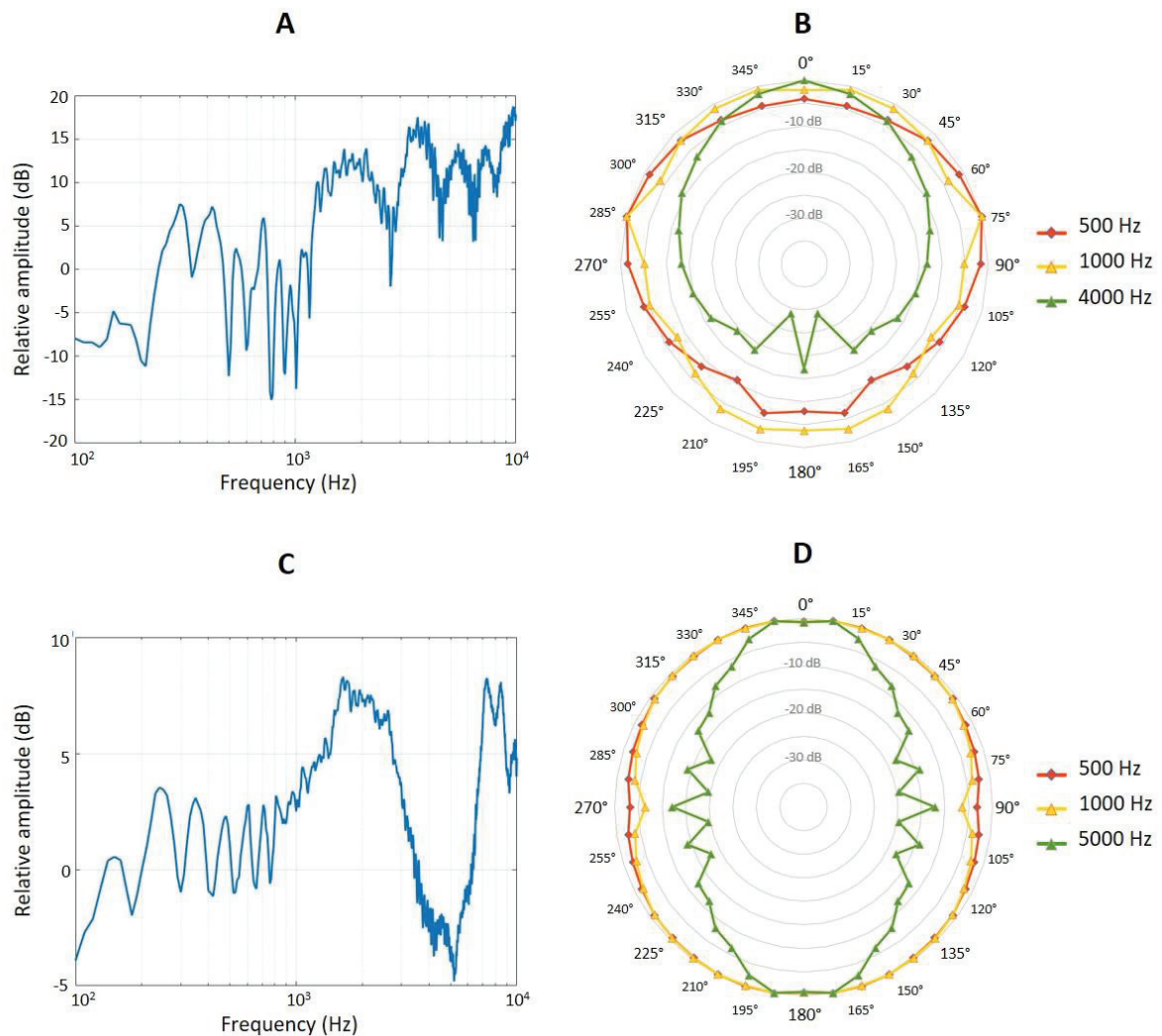
Our final experiment questions the acoustic basis for the categorization performed by crocodilians. It turns out that crocodilians rely on the spectral envelope of sound signals to categorize stimuli. However, they ignore the pitch and aperiodic component of the signals. Using the spectral envelope to identify a complex sound made up of a series of harmonics is a widespread and reliable strategy (Burgering et al., 2019). The spectral envelope, corresponding to the timbre of the emitter, is an invariant acoustic parameter, while the pitch evolves for a same individual with the age (A. Vergne et al., 2007) and the context, the spectral envelope would be supposed to be a more stable and reliable cue to categorize crocodile calls, all types of calls combined.

In conclusion, our study shows that crocodilians exhibit an ability to categorize sound stimuli. This ability certainly contributes to an efficient apprehension of auditory scenes in their everyday life. It complements other processes such as their ability to locate precisely the position of a sound source (Bierman & Carr, 2015; Carr & Christensen-Dalsgaard, 2015; Papet et al., 2019, 2020) and their ability to identify a meaningful sound despite background noise by spatial unmasking (Thévenet et al., 2022b). Given their lifestyle, it is important that these top predators possess powerful abilities to separate and identify sound sources, and more generally to apprehend their auditory sensory space. Crocodilians are indeed confronted with complex sound environments, and must be able to make quick decisions. Being able to quickly categorize sounds according to their functional relevance is certainly extremely useful.

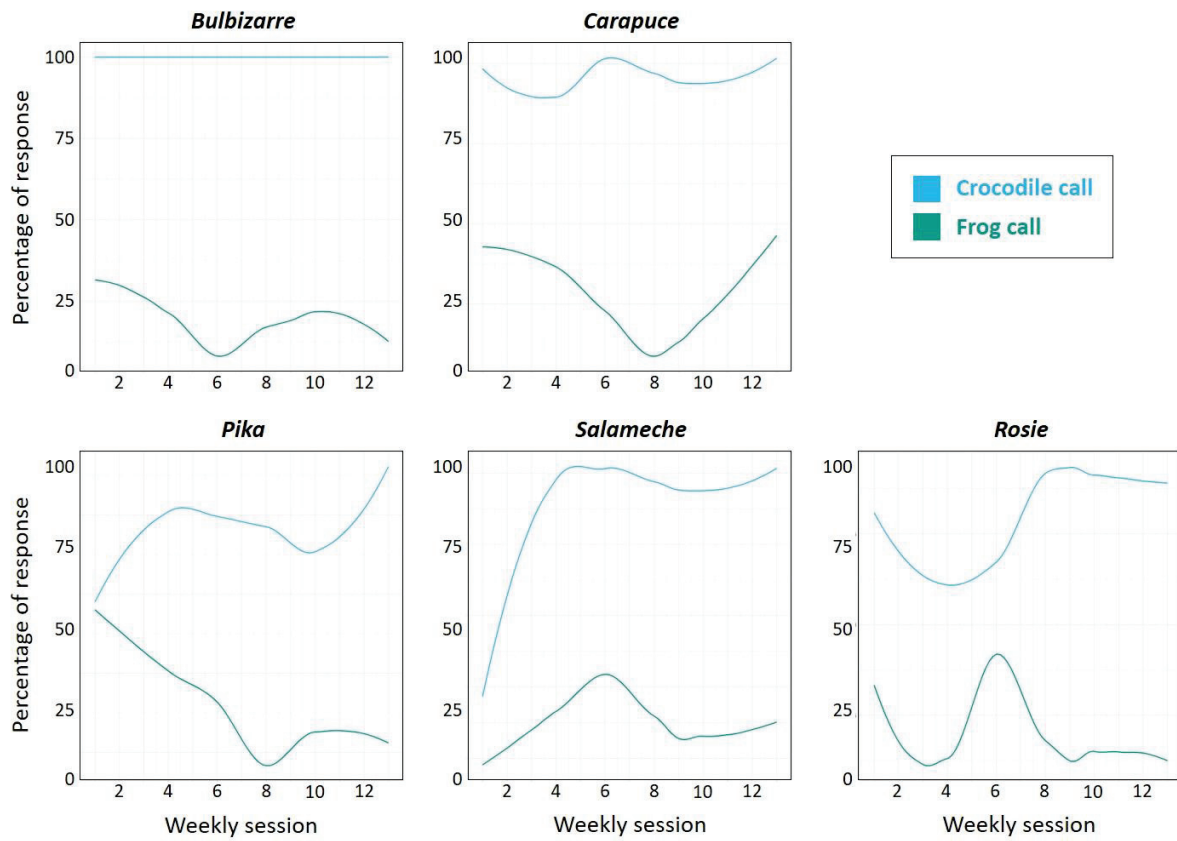




## 1.5 Supplementary information



**Supplementary Figure IV.1. Loudspeakers technical specifications.** All measurements were done in a semi-anechoic chamber (3.4m x 4.6m, reverberation time = 0.2s for 125 Hz and  $\leq$  0.1s for frequencies higher than or equal to 500 Hz). **A.** Frequency response of the FoxPro Fusion speakers (internal speakers Visaton #SL 87 ND) used in Experiment 1. **B.** Directivity pattern of the FoxPro Fusion speakers used in Experiment 1. **C.** Frequency response of the AudioPro, Bravo Allroom Sat speakers used in laboratory experiments (experiment 2 and 3). **D.** Directivity pattern of the AudioPro, Bravo Allroom Sat speakers used in laboratory experiments (experiment 2 and 3).



Supplementary Figure IV.2. Learning curves of the five juveniles Cuvier's dwarf caimans (*Paleosuchus palpebrosus*).



## 2. Pupillometry experiments

We propose here to pursue this study on sound categorization by young crocodylians through an innovative approach that is pupillometry. I started to implement this methodology in order to complement our auditory discrimination experiments and thus to answer more rigorously to the question of the category boundary effect.

### 2.1 Pupillometry principle

Pupillometry is of growing interest for the study of physiological and perceptual phenomena. Besides the well-known pupillary light reflex phenomena, in which the pupil constricts to narrow the light intensity reaching the retina under high level of luminance (Ellis, 1981; Schaeffel & Wagner, 1992; Li & Howland, 1999; Barbur et al., 2002; Douglas et al., 2002, 2005), the pupil size also varies with the internal physiological state of the subject, related to mental effort or change in arousal level (Reinhard & Lachnit, 2002; Mathôt, 2018).

Under constant light condition, the pupillary diameter is indeed under the control of the autonomic (sympathetic) nervous system. While the precise neural circuit is still not fully understood, variations in pupil size have been associated to Locus Coeruleus (LC) and the norepinephrine (NE) activity (Joshi et al., 2016). When organisms face up a stress, the activation of radial muscles by sympathetic neurons lead to pupil dilatation while the constriction is caused by a stimulation of circular muscles by parasympathetic neurons.

The activation of the norepinephrine system and change in arousal level is commonly studied by recording the magnitude of brain-evoked potential components (see Fabiani et al., 2007; for a review). In particular, the P3 ERP component has been associated with concomitant changes in pupil size as a result of modulation of arousal (Nieuwenhuis et al., 2011). As such, the pupillometry methodology offers great promise for studying cognitive processes in non-human animals in a non-invasive way, that is more or less as effective as EEG.

Numerous cognitive processes has been studied using pupil size indicator in humans, including attention (Wierda et al., 2012; Kang et al., 2014; Brink et al., 2016), surprise effect (Preuschoff et al., 2011; Lavin et al., 2014; Kloosterman et al., 2015; Alamia et al., 2019), decision-making process (Cavanagh et al., 2014; de Gee et al., 2014; Murphy et al., 2014), uncertainty (Richer & Beatty, 1987; Lavin et al., 2014; Urai et al., 2017), emotion (Bradley et al., 2008).

In particular, variations in pupil size have been shown to be modulated by *oddball* effects, which occur when stimuli have uneven probabilities of occurrence (emergence of rare stimuli compared to frequent ones). Pupil dilation is inversely related to stimulus probability, i.e., rarer stimuli cause greater pupil dilation (Reinhard & Lachnit, 2002; Gilzenrat et al., 2010). This probability effect has been supposed to be driven by centre processes such as expectations or surprise, linked to bottom-up effects of attention drawing, more than by the nature of the stimulus itself (Friedman et al., 1973; Qiyuan et al., 1985). In the auditory oddball task, rare stimuli (“deviants”) are embedded in a train of frequent stimuli (“standards”). The subject is habituated to the standard stimulus that is repeated in a predictable way, and will be thus surprised by the unpredictable emergence of deviant stimuli (Bala & Takahashi, 2000; Montes-Lourido et al., 2021). If the deviant stimulus is perceived as different as the standard one, it should elicit an internal reaction of the subject such as a change of pupil size.

Investigating precise auditory discrimination threshold in crocodilians is a complicated task (cf. discussion *part IV.1.4*). To go further in our study of categorical perception of sound stimuli by crocodiles, I therefore proposed to use an auditory oddball paradigm with a pupil dilation measurement. I made the hypothesis that sound stimuli belonging to different perceptual categories would be better discriminated and would elicit a greater pupil dilation when paired as standards and deviants in the same sequence.

## 2.2 Particularity of the crocodilian’s eye

Before talking about the protocol in more detail, just a few words about the eye anatomy of crocodilians. The eyes of crocodilians have a basic vertebrate anatomy: a curved cornea covers the iris and lens. The light passes through the lens, the vitreous body and reaches the retina to be absorbed by the pigments of the photoreceptors. The signal is transduced electrically via the ganglion cells and the optic nerve to the brain. The retina contains one type of rod cells, that are the most sensitive under low-light conditions, and cones (which composition depends on the species, see Reber, 2020 for a review), that enable color visions (light absorption between 424 nm – violet, and 566 nm – yellow). Rods are the most numerous photoreceptor in the crocodilians retina, suggesting that the retina may have evolved to improve vision at night (Sillman et al., 1991). An important aspect to be noted is the lack of evidence to date that crocodilians can perceive ultraviolet light (Nagloo et al., 2016). As for nocturnal birds and



mammals, crocodylians also have a tapetum lucidum, a reflective layer behind the retina acting as a light amplifier. The neurophysiological pathway between the retina and the crocodylians' brain has been found to be very similar to that of the birds, suggesting a relatively well-developed binocular vision (Shimizu et al., 2009; Grigg & Kirshner, 2015). Their field of vision is relatively large, supposedly up to 260 degrees (Grigg & Kirshner, 2015)

Crocodylians have this characteristic of having two eyelids and a nictitating membrane closing the eye when the animal immerses. More important for the implementation of the pupil measurement protocol, the pupil is a vertical slit when exposed to intense light and becomes rounder as the light decreases (Grigg & Kirshner, 2015), which can greatly complicate the tracking. Besides the fact that the pupil is an oval slit, another major difficulty encountered with these animals was the necessity to restrain them in order to immobilise their head. Dwarf caimans have this typical head-up posture that made it necessary to adapt a restraint system that is both non-invasive and as comfortable as possible for them (involving the creation of a trough in which the animal was placed with an elevated headpiece). Unfortunately, the lack of time and means due to the pandemic, and the complexity to apply such a protocol with crocodiles slowed down and limited the acquisition of data. I thus present here only preliminary results.

## **2.3 Methodology**

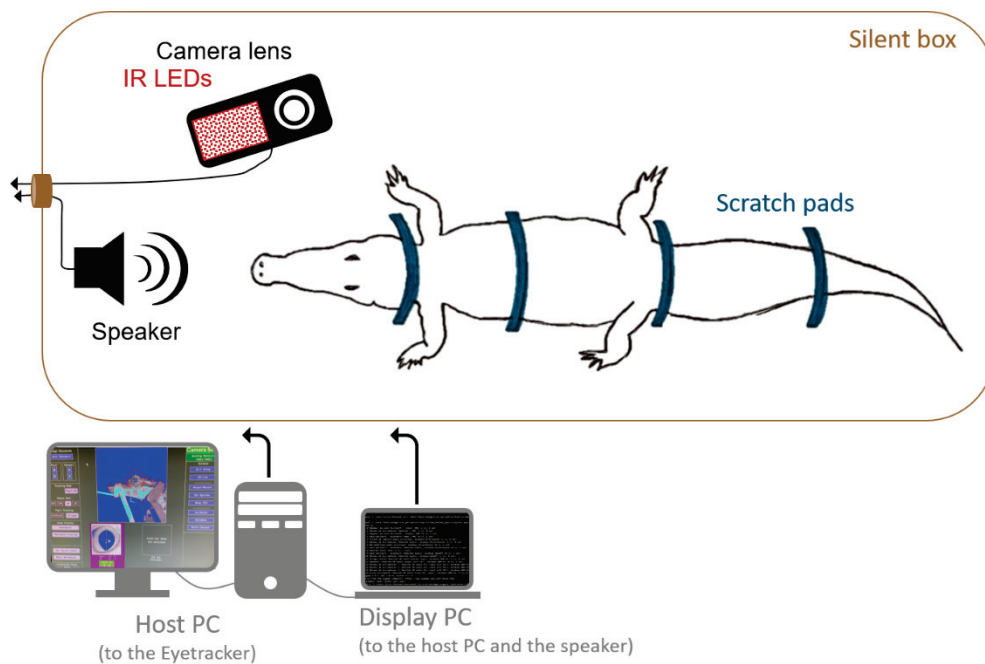
### **2.3.1 Experimental setup**

We worked with the same 5 dwarf caimans as the previous experiments on categorization. The animals were immobilized with a custom-made system, placed in the centre of a soundproof booth (Figure IV.4). The pupil size of the right eye was monitored using an eye tracker EyeLink 1000 Plus (SR Research), placed at 40 cm from the eye. The tracking system was set with a EyeLink 1000 plus High Speed 35mm Desktop Lens, with a sample rate of 1kHz and a medium illumination power of 75%. Pupil size was calculated as an area. The pupillometry data were retrieved live from the experimenter's computer, which communicated with the host PC using a Python script written by Alexandra Corneyllie.

The illumination level was kept constant during all the duration of the experiment, to avoid bias due to pupillary light reflex. The Eyetracking is a system originally designed for pupil analysis in humans. So, I faced up the challenge to find the right parameters to adapt this system to the analysis of a pupil with a vertical slit. In particular, the light intensity was set so that the system

would track the pupil as accurately as possible: not too much light to limit the thinning of the pupil which becomes too oval for the system to track, and not too much darkness to avoid having a pupil which is too dilated and which makes it difficult for the system to track because it reduces the contrast area between the iris and the pupil.

The stimuli were presented from a loudspeaker placed at 40 cm in front of the animal's head (to prevent head's movement due to a sound source location). The entire testing sessions were also live-filmed and recorded with an infra-red camera (ABUS TVCC34010) and controlled by the experimenter from outside the booth.



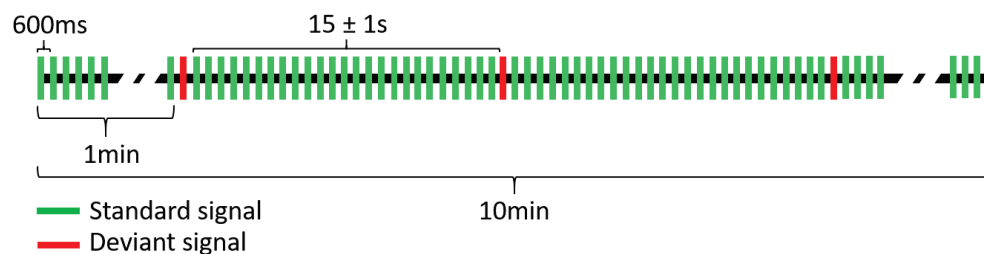
**Figure IV.4.** Pupillometry setup. The host computer is linked to the camera and manages the tracking system. The display computer is used to control the sound sequence emission and to recover pupillometric data from the host PC.

### 2.3.2 Acoustic stimuli

Two types of auditory sequences were synthesised (Matlab script written by Naïs Caron Delbosc). One sequence was constructed with the 80% frog / 20% crocodile call as the standard stimulus, the 100% frog / 0% crocodile and 60% frog / 40% crocodile calls as the deviants, and 80% frog / 20% crocodile as a control deviant. The other type of sequence was constructed with the 60% frog / 40% crocodile call as the standard stimulus, the 80% frog / 20% crocodile and 40% frog / 60% crocodile calls as the deviants, and the the 60% frog / 40% crocodile as a



control deviant (Figure IV.5). These two sequences allowed me to test the discrimination of both stimuli belonging to the same category and stimuli from different categories at the same time. Within a sequence, the stimuli were emitted at a rate of one call every 600ms over 10 minutes. The sequences always began with at least 1 minute of standard stimuli to habituate the pupil response. During the 9 following minutes, deviant stimuli were then presented < 10% of the time, with a minimal interval between two deviants of 15sec to let the pupil response reach its maxima and then return to normal. The position of deviant stimuli in the sequences was pseudorandomized: the three types of deviant (first deviant, second deviant, control deviant) were always placed one after the other in a random order, and this randomly drawn triad was repeated during the 9 minutes of the sequence. A unique copy of sequence was generated for each caiman, so that each caiman is tested only once with each of the 2 types of sequence.



**Figure IV.5.** Schematic description of a sound sequence used in the pupillometry experiments.

### 2.3.3 Experimental procedure

Prior to the testing period, the caimans were acclimatized to being immobilized with a custom-made system, placed in the centre of a soundproof booth for progressively longer period of time, over two months. When the animals were accustomed to the restraint system, they were each tested in a single session. A test session began with immobilizing the caiman and left him without any stimulation for an acclimation period of at least 10 min. The pupil diameter was continuously monitored from the onset of the acclimation period to 15 sec after the onset of the last stimulus. A testing session was interrupted for one minute if the animal struggled and restarted again with a new sound sequence, or was aborted if the caiman struggled more than twice.

### 2.3.4 Analysis

All analysis was performed in Matlab using a dedicated script written by Naïs Caron Delbsoc. The raw pupil size data acquired by the eyetracker were first linearly interpolated, since the

eyetracker did not have a fixed sampling rate during the experiment (mean  $870 \pm 151$  Hz). The spectrum of the signal obtained was then calculated in order to create 2 filters: a low-pass filter (Butterworth type) with a cut-off frequency of 3.5 Hz, and a high-pass filter (Butterworth type) at 0.001 Hz to cut the continuous component.

The position of each deviant within the signal was found thanks to the position files created during the synthesis of the sequences giving the temporal index of each. While viewing the signal trace, we manually removed the deviants for which we observed jumps or losses in pupil size. As a preliminary experiment, we relied on the literature (see Montes-Lourido et al., 2021) as well as our own observations to choose a temporal observation window. We chose a wide window ranging from 1 second before and 4 seconds after the emission of each deviant. Since the goal was to observe the variation in pupil size upon deviant emission, the pupil size baseline was corrected to zero between  $t-1$  and  $t_0$ .

Finally, for each of the three types of deviant for each standard (20 or 40) and for each crocodile, the average of all the vectors of pupil size between  $t-1$  second and  $t+4$  seconds with respect to  $t_0$  the emission of the deviant is calculated. When for the same caiman two sequences of the same standard could be tested (one occurrence), the global average on these two sequences was calculated in order to obtain an average result per caiman and per type of deviant. Finally, we also calculate the average of each type of deviant for the two types of sequences (standard 20 or 40) for all the caimans combined.

## 2.4 Preliminary results

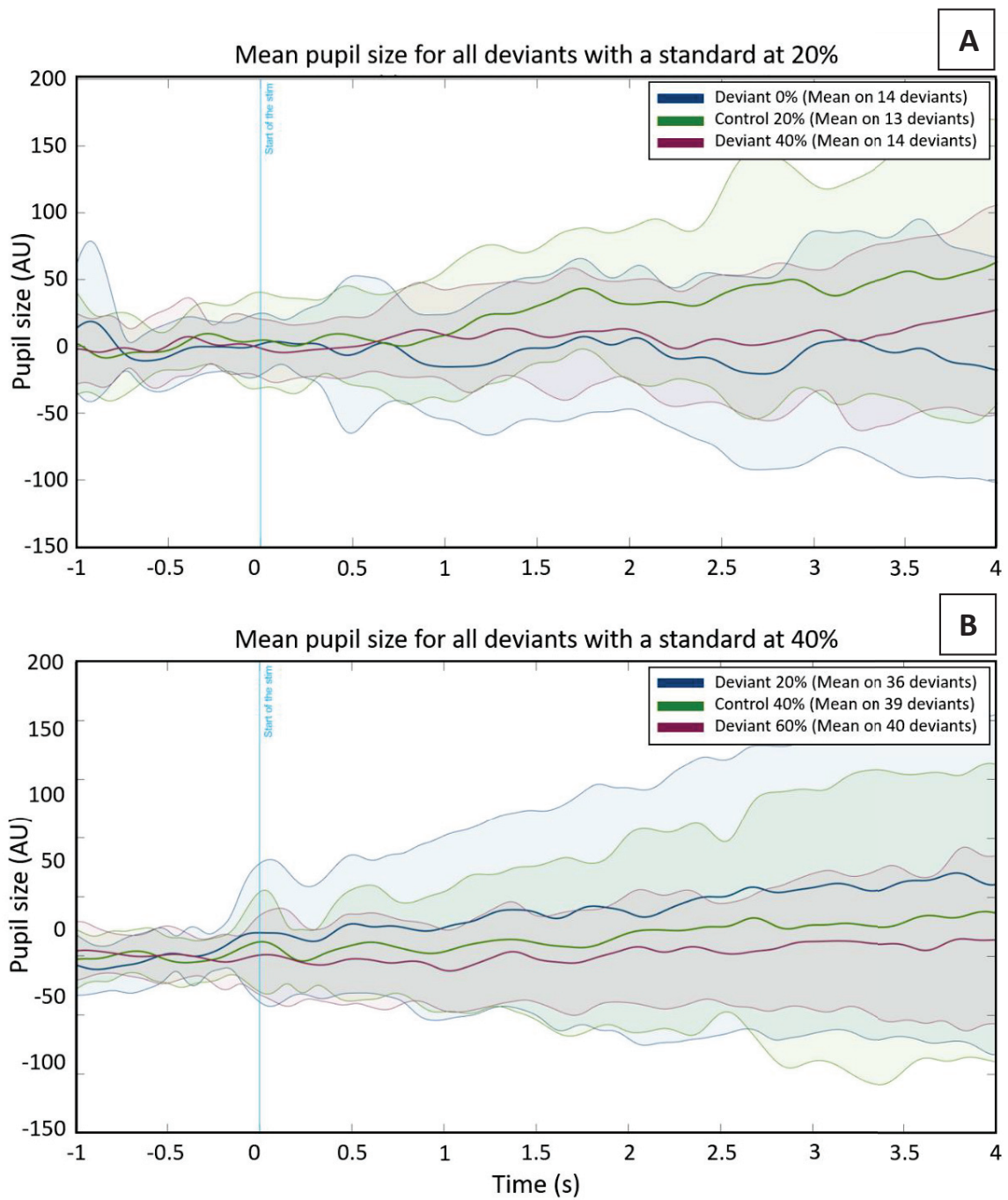
Among the 5 caimans tested, only 2 could be tested with both types of sequences (standard 20% and standard 40%), one of which could be tested twice with the same sequence (resulting in only 8 usable result files). Of these files, where 30 deviant signals were emitted over 10 minutes, only [21, 15, 20, 14, 6, 25, 28, 25] responses to the deviant signals were usable. The rest of the time, either the crocodile was moving or the system did not reliably track the pupil size.

I present here the averaged results for all caimans for the sound sequences based on the standard signal 80% frog / 20% crocodile call, and for the sequences with the standard signal 60% frog / 40% crocodile call (Figure IV.6).

In the first sequence (standard with 20% crocodile call; Figure IV.6A), we hypothesized that the change in pupillary response would be greater for the 40% crocodile call deviant than the







**Figure IV.6. Average pupil size for sequences based on the standard signal corresponding to the 80% frog / 20% crocodile call (A) and for the sequences with the standard signal 60% frog / 40% crocodile call (B). The green curves correspond to control stimulus (same signal as the standard. Shadows are the standard deviation calculated on all deviants on all subjects.**

0% crocodile call and the control, since the perceptual boundary previously established was between the 20% and 40% crocodile calls and that we should have a peak of discrimination at the category boundary. Here, the pupillary response to the 40% deviant signal appears to be different from that of the control signal 200 ms after stimulus emission. However, the pupillary response to the 0% deviant signal also appears to be different from that of the control. Thus, the two deviant signals appear to elicit an equivalent pupillary response.

When the animals were tested with the standard sequence consisted of the 40% crocodile call (Figure IV.6B), the results are a little different since it seems that the pupillary response curve to the 20% crocodile call deviants seems to be slightly different from that of the controls and of the 60% crocodile call deviants, whereas the curve of the 60% crocodile call deviants deviates only slightly from that of the control. However, the standard deviations of the pupillary response curves are too high for these results to be conclusive.

## 2.5 Perspectives

Finally, even if we observe a tendency, our results obtained so far do not allow us to confirm our hypotheses: the pupillary response to the different types of deviants does not seem to follow a pattern allowing us to conclude that this measure is an index of auditory discrimination.

Unfortunately, to date, only few data could be acquired for these experiments for 2 main reasons. The first one being that the dwarf caimans did not support the restraint, maybe partly because of their nervous character and their typical head-up posture which makes the immobilization of the head very uncomfortable for them. In fact, many test attempts had to be stopped because the animal was moving too much. The second major problem, which I expected, is the tracking system of the Eyelink, which is basically configured for the human eye. I had to deal with a lot of data loss, as the eyetracker could not track the pupil correctly at many points.

In sum, the analysis of pupil response in crocodylians remains a promising tool for the study of physiological responses, whether auditory, olfactory or visual. To date, the data are yet not conclusive and it is necessary to develop a system adapted to the anatomy of crocodiles, both in terms of tracking and restraint. The choice of the species to be studied is also a major criterion to consider. More docile species with their head naturally flattened against the ground, such as the alligator, will be more suitable for this type of study.





# V. Hearing others: perceiving heterospecific distress

## Crocodile perception of distress in Hominid baby cries

This article is in revision for the *Proceedings of the Royal Society B* journal.

**Julie Thévenet<sup>1,2</sup>, Léo Papet<sup>1,2</sup>, Gérard Coureaud<sup>1</sup>, Nicolas Boyer<sup>1</sup>, Florence Levréro<sup>1</sup>,  
Nicolas Grimault<sup>2,‡</sup>, Nicolas Mathevon<sup>1,3,‡</sup>**

<sup>‡</sup> Co-last authors

<sup>1</sup>Equipe Neuro-Ethologie Sensorielle ENES / CRNL, CNRS, Inserm, University of Saint-Etienne, France;

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

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

## Abstract

Is there a universal acoustic coding of emotion? It is generally argued that distress vocalizations, a common modality for alerting conspecifics across a wide range of terrestrial vertebrates, share universal acoustic features that allow heterospecific understanding. Yet studies suggest that the acoustic traits used to decode distress may vary between species, leading to decoding errors. Here we show through playback experiments that crocodiles are attracted to infant hominid cries (bonobo, chimpanzee and human), and that the intensity of crocodile response depends critically on a set of specific acoustic features (mainly deterministic chaos, harmonicity, and spectral prominences). Our results demonstrate that crocodiles are sensitive to the degree of distress encoded in the vocalizations of phylogenetically very distant vertebrates. A comparison of these results with those obtained with human subjects confronted with the same stimuli further shows that crocodiles and humans use different acoustic criteria to assess the distress encoded in infant cries. Interestingly, the acoustic features driving crocodile reaction are likely to be better universal markers of distress than those used by humans. These results highlight the universality of acoustic features encoding distress in vertebrate vocalizations, while demonstrating that the way they are used to decode distress may vary across species.



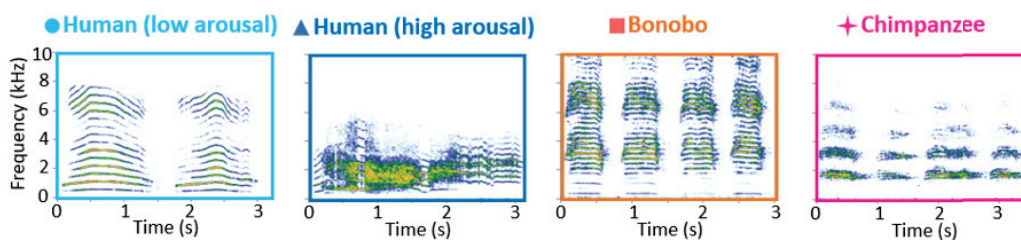
# Graphical abstract

## Crocodile perception of distress in Hominid baby cries

Female crocodilians are very sensitive to the distress calls of juvenile crocodiles. As top predators with good auditory abilities, are they also able to perceive distress encoded in the vocalizations of other species?

It is generally argued that distress vocalizations, a common modality for alerting conspecifics across a wide range of terrestrial vertebrates, share universal acoustic features that allow heterospecific understanding. Yet studies suggest that the acoustic traits used to decode distress may vary between species, leading to decoding errors. A scientific approach, based on acoustic analyses of emotion-carrying signals and on playback experiments involving phylogenetically distant species is needed to identify the acoustic parameters relevant to the receivers.

### Playback experiments

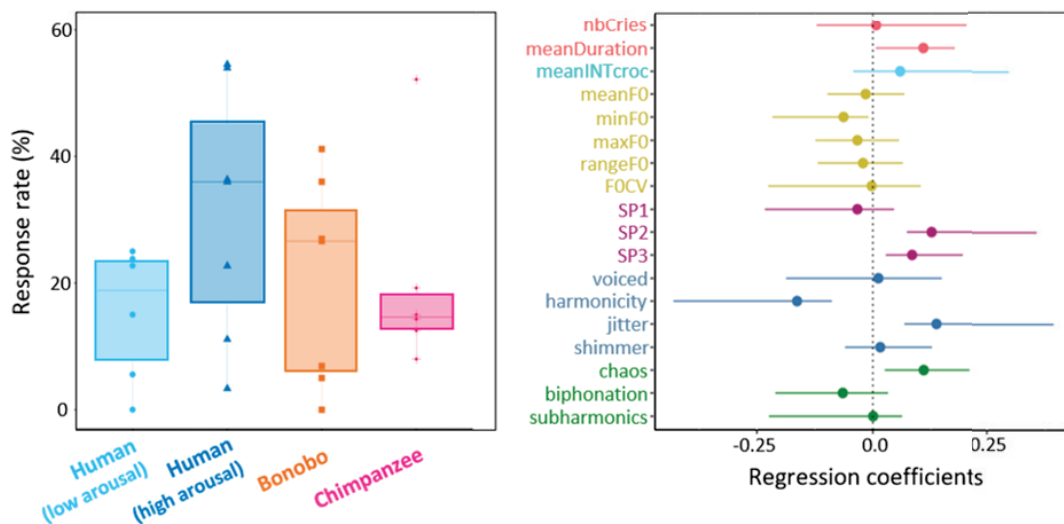


Playback to groups of Nile crocodiles



Analyze of the rate of responding animals within a group depending of the sound signal

### Results



The intensity of crocodile response depended on a set of specific acoustic features (mainly deterministic chaos, harmonicity, and spectral prominences) rather than to a particular type of call

### Conclusion

Crocodiles are sensitive to the degree of distress encoded in the vocalizations of phylogenetically distant vertebrates. The acoustic features that motivate crocodile response are likely universal markers of distress encoded in vertebrate vocalizations. However, the acoustic features that elicit their attention are different from those used by humans, demonstrating that the way they are used to decode distress may still vary between species.

# 1. Introduction

The cross-species perception of emotions conveyed by voice is a topic that tickles both biologists and the general public, as it challenges our ability to understand what non-human animals convey through their vocalizations (Briefer, 2020; Filippi et al., 2017; Scheumann et al., 2014). People who live with pets often claim to be able to decipher emotions such as distress, joy or surprise, by listening to their companion meow or bark (Prato-Previde et al., 2020; Faragó et al., 2017, 2014). Livestock and pet professionals often agree (Laurijs et al., 2021; Manteuffel et al., 2004; Tallet et al., 2010). However, individual impressions and empirical observations, while suggesting the existence of inter-species understanding, do not provide objective information about the decoding processes of the acoustic emotional signals involved. A scientific approach, based on acoustic analyses of emotion-carrying signals and on playback experiments to identify the acoustic parameters relevant to the receivers, is needed to shed light on the processes involved.

Darwin had already hypothesized that natural selection had led to convergences in emotion coding between animal species, and that the expression of emotions by voice has ancient evolutionary roots. He suggested that the way emotions are expressed vocally must be sought in the mechanisms of production by the vocal organs (Darwin, 1871). Indeed, since most airborne vertebrates share the same principles of sound production (vibration of membranes driven by airflow), the acoustic output of emotions is likely to be similar in all these animals, suggesting that they therefore share similar acoustic coding of emotions. A century after Darwin, Morton (Morton, 1977) stated the principle that the acoustic structure of a sound signal should reflect the motivational state of the sender (motivation-structural rules): birds and mammals use rather low-pitched and "rough" sounds in hostile situations, and rather high-pitched and pure tone sounds when they are frightened, or animated by friendly motivation. Emotions can be classified along two dimensions: their arousal level (low to high) and their valence (negative or positive) (Laukka et al., 2005; Mendl et al., 2010). Recent studies provided empirical data supporting Morton's claim that vocalizations are usually louder and produced at a faster rate, higher pitched, more frequency-modulated and rougher when arousal increases, with positive vocalizations being shorter and less frequency-modulated than negative vocalizations (August & Anderson, 1987; Briefer, 2012; Ehret, 2006). Emphasis has been placed on nonlinear phenomena, i.e. irregularities in acoustic signals that arise from perturbations of vocal folds vibrations (Fitch et al., 2002). These nonlinear phenomena (Wilden



et al., 1998) (frequency jumps, subharmonics, deterministic chaos and sidebands) are the main cause for the typical roughness of vocalizations related to high arousal, such as distress calls (Lingle et al., 2012; Massenot et al., 2022; Anikin et al., 2021; Marx et al., 2021; Blumstein et al., 2008; Stoeger et al., 2011). Nonlinear acoustic phenomena are therefore good candidates for a universal coding of distress across animal species.

If commonly shared, these rules of coding lead to the hypothesis that airborne vertebrates - including humans- are able to decode emotions conveyed by heterospecific acoustic signals on the basis of universal acoustic features. Although a number of studies have supported this hypothesis by showing, for example, that human subjects are able to estimate the degree of arousal in the calls of different vertebrate species (Faragó et al., 2014; Filippi et al., 2017; Koutseff et al., 2018; Maruščáková et al., 2015; McComb et al., 2009; Pongrácz et al., 2006), other studies suggest however that the acoustic criteria used by receivers are not always appropriate to correctly assess the level of distress. For example, Kelly et al. (Kelly et al., 2017) showed that human listeners assess the level of distress encoded in ape infant cries (human, bonobo and chimpanzee) based on pitch, leading them to overestimate the level of distress encoded in all the cries of bonobo infants which are very high-pitched. Similarly, Teichrob et al. (Teichrob et al., 2013) and Lingle and Riede (Lingle & Riede, 2014) find that female mule deer and white-tailed deer respond to distress calls from different mammals only when the frequency range of these calls is artificially brought into the frequency range of baby deer vocalizations. Root-Gutteridge et al. (Root-Gutteridge et al., 2021) also showed that the response of domestic dogs to the cries of puppies and human babies is highly dependent on their frequency range. Thus, the ability to identify emotions across species may be limited by species-specific traits (Bremond, 1976). Furthermore, we have very limited knowledge of how non-human animals can decode emotions in the vocalizations of other animal species since most studies focus on the perception of animal vocalizations by humans. Moreover, investigations in non-human animals usually involved species that are phylogenetically close (e.g., cross-species perception of alarm and distress calls in birds (Aubin, 1991); deer perception of distress calls from other mammals (Lingle & Riede, 2014)). To further explore the hypothesis of universal emotion coding in vocalizations and to better understand the mechanisms, it could be necessary to conduct studies where individuals are tested with acoustic signals from phylogenetically distant species.

In the present study, we investigated the perception of distress encoded in the cries of Hominid infants (bonobos, chimpanzees and humans) by Nile crocodiles. Crocodylians and hominids are

distant vertebrate groups on the phylogenetic tree. Birds and mammals are almost the only non-human animals that have been studied for heterospecific communication and tested for human voice perception. Crocodylians are the closest living relatives of birds and possess several ecological characteristics that have evolved in a convergent manner with mammalian species, although they are very distant on the phylogenetic tree (Nesbitt, 2011; Grigg & Kirshner, 2015; Stevenson, 2019). In particular, both crocodylians and hominid infants have this characteristic to solicit parental care through distress vocalizations (A. Vergne et al., 2009; Bard, 2000; Bermejo & Omedes, 2000) (hereafter called “cries”). Caregivers -a parent in most species- are attracted to these infant cries and respond by providing various types of parental care (protection from predators in crocodiles; protection and feeding in hominids). Moreover, crocodiles are top predators and sometimes commit cannibalism (Hutton, 1989; Grigg & Kirshner, 2015). They may be attracted to the cries of potential prey. Although their auditory and sound localization performance has been well studied (Carr & Christensen-Dalsgaard, 2015; Papet et al., 2020; Reber, 2020), their perception of their sound environment and their ability to extract and use information from it remains understudied to date. Crocodiles are therefore an excellent model for investigating an animal's ability to identify distress in the vocalizations of other phylogenetically distant species. Here we test the hypothesis that crocodylians react to the distress encoded in the cries of infant hominids, and that the amplitude of their reaction is driven by the magnitude of specific acoustic features. We first performed an analysis pointing out the differences of acoustic structures between the cries of hominid babies (bonobos, chimpanzees and humans) recorded in different situations eliciting various levels of distress. We then observed the behavioral response of adult Nile crocodiles to the cries and identified the acoustic characteristics of the cries (all species combined) that explain the variation in reaction intensity of the crocodiles. Finally, we compared these results with those obtained in adult humans in a previous study using the same stimuli (Kelly et al., 2017). We show that, unlike adult humans whose response to baby cries is essentially driven by cry frequency pitch, crocodiles are particularly attentive to a set of features including nonlinear acoustic phenomena, making these animals likely to be more effective at identifying levels of distress encoded in vocalizations from phylogenetically distant species.





## 2. Methods

### 2.1 Acoustic stimuli

All stimuli have been recorded prior to the present study and are part of the sound database of the ENES laboratory (Figure V.1A). They were selected for recording quality and to cover a wide distribution of acoustic parameters.

The context of bonobo and chimpanzee infant cries' recordings were either begging the mother for assistance while no danger was apparent or calling for being protected during aggressive interactions. Given the diversity of recording situations, we assume that our recordings represent a fair distribution of arousal levels coded by bonobo and chimpanzee infant cries. Bonobo infant cries were recorded in European zoological parks, with a Sennheiser MKH70 ultra-directional microphone connected to a Zoom H4n recorder (sampling frequency = 44.1 kHz). The recorded infants were aged from 1 to 4 years ( $N = 6$ ) and all dependent on their mother (breast-fed and frequently carried by the mother). We isolated 6 cries from the recordings (cry duration =  $3.0 \pm 0.4$  seconds). Chimpanzee infant cries were recorded in the wild from a population habituated to humans (Kibale National Park, Uganda), with Sennheiser MKH70 ultra-directional microphone connected to a Marantz PMD670 digital recorder (sampling frequency = 44.1 kHz). The exact ages of the recorded infants ( $N = 7$ ) were unknown, but all were under 4 years and were carried by their mother. We isolated 6 cries from the recordings (cry duration =  $3.2 \pm 0.4$  seconds).

Human babies' cries were recorded in two contexts: bathing at home by parents (low arousal context;  $N = 6$  babies with balanced sex ratio) and vaccination at the pediatrician's office (high arousal context;  $N = 6$  babies with balanced sex ratio, different from those recorded in the bathing context). The babies were recorded with a Sennheiser MD42 microphone placed at 30 cm from their mouth, and connected to a Zoom H4n recorder (sampling frequency = 44.1 kHz). We isolated one cry per recorded baby (total of 12 sequences; duration of each cry =  $3.2 \pm 0.2$  seconds).

### 2.2 Analysis of sound stimuli

We analysed the acoustic structure of cry sequences using a custom script in *Praat* software (Boersma & Weenink, 1992). Extending the analysis method developed by Kelly et al. (2017),

we measured the following eighteen acoustic variables: number of cry syllables in the sequence (*nbCries*), average duration of each syllable in the sequence (*meanDur*), percentage of the sequence duration with detectable pitch (*voiced*), average pitch over the sequence (*meanF0*), minimum pitch (*minF0*), maximum pitch (*maxF0*), range of pitch ( $\text{maxF0} - \text{minF0} = \text{rangeF0}$ ), pitch coefficient of variation (*F0CV*), harmonicity (*harmonicity*; ratio of harmonics to noise in the signal expressed in dB), jitter index (*jitter*; small fluctuations of periodicity), index of shimmer (*shimmer*; small variations in amplitude), first three spectral prominences characterizing the spectral envelope of the cries (*SP1*, *SP2*, *SP3*), percentage of the sequence duration with subharmonics (*subharmonics*; nonlinear phenomena appearing on the spectrogram as integer fractional values of an F0), percentage of the sequence duration with biphonation (*biphonation*; nonlinear phenomena characterized by two simultaneous and independent fundamental frequencies), percentage of the sequence duration with deterministic chaos (*chaos*; nonlinear phenomena characterized by non-random noise), mean intensity of the cry sequence (*meanINTcroc* or *meanINThuman*). The latter variable was calculated by considering the respective auditory sensitivity of crocodiles and humans. The maximum amplitude of all signals was previously normalized. Based on the crocodilian audiogram measured by Higgs et al. (2002), we converted flat dB into “dB crocodile” and obtain the average intensity corresponding to the hearing sensitivity of crocodiles (see Supplementary Material 1 for details). The average human intensity was calculated using dB(A).

To illustrate the differences between cry categories while reducing the acoustic dimensions considered, we performed a principal component analysis (PCA) on all 18 acoustic variables (taking *meanINTcroc* as the measure for the mean intensity of the cry sequence). Clustering of cry categories according to the first three principal components (acoustic dimensions) was tested using anova (package *stats*, R-Studio v.4.1.2), taking each acoustic dimension as the dependent variable and the cry category (bonobo, chimpanzee, human baby at bath, human baby during vaccination) as the fixed factor. We then performed post hoc multiple comparisons of means (Tukey contrasts, R package *multcomp*).

## 2.3 Playback experiments on crocodiles

The experiments were conducted at CrocoParc zoo (Agadir, Morocco). This park hosts more than 300 Nile crocodiles (*Crocodylus niloticus*) in an outdoor garden including several ponds. As the animals are free to roam, it was impossible to test the crocodiles individually. We



therefore played back the sound stimuli to four groups of between 7 and 25 adults (females and males, unknown sex ratio), occupying 4 different ponds (see appendix Supplementary Figure V.1 for a plan of the ponds and the position of the speakers).

In order to avoid habituation, each group of crocodiles was tested during a single experimental session, except for one group that was tested in two sessions, one day apart. To accustom the crocodiles to the presence of the speakers, they were positioned two days before the start of the experiments. This ensure us a control condition where no sounds were played and where we could verify that the crocodiles do not come towards the silent speakers (our previous experiments with these same animals also allow us to affirm that they do not respond to any type of signal like white noise (Thévenet et al., 2022b)). Each experimental session then started at 19:00, one hour after the park closed to the public.

During each experimental session, each group of crocodiles heard a succession of up to 7 stimuli (min-max = 5-7), broadcasted with a remote-controlled loudspeaker (FoxPro Fusion with Visaton SL 87 ND internal speakers, see appendix Supplementary Figure V.2 for the technical specifications). Each stimulus consisted of a 30-second repetition of one of the 24 previously isolated cries. The stimuli were different between groups. Among the stimuli sent to each group, there was at least one cry from each category (bonobo, chimpanzee, human baby bathing, human baby being vaccinated; see appendix Supplementary Table V.1 for a detailed list of stimuli played to the different groups), with a minimum interval of 10 min between each playback.

All experiments were filmed (Lumix DMC-FZ300 camera). The behavioural response of the crocodiles was assessed by measuring the proportion of individuals who responded to the stimulus (number of individuals who turned their head toward the speaker or moved in its direction divided by the number of individuals present during the experiment).

## **2.4 Analysis of the crocodile reaction to playback**

We tested the effect of the cry category on crocodile behavioral response with a generalized mixed model (package *lme4*, logistic function, R-Studio v.4.1.2, fixed factor: cry category, random factor: number of the pond where the tested group of crocodiles was located). We used multiple comparison tests to compare the intensity of crocodile behavioral responses across stimulus categories.

To identify the acoustic traits that could explain crocodiles behavioral responses, we explored the relative importance of each of the 18 acoustic variables characterizing the stimuli using partial least squares logistic regression PLS (package *plsRglm* (Bertrand et al., 2009)). PLS is useful when a response has to be predicted from a large set of variables and when there is multicollinearity. While classical principal component analysis does not identify the salient acoustic features explaining the behavioral response, PLS-regression allows to group the acoustic features that best predict these responses. PLS constructs components from linear combinations of the predictors optimised to be related to the variable to be explained. Here, the variable to be explained was the crocodile behavioral response while the predictor were the 18 acoustic parameters. Cross validation was used to select the optimal number of components in the model. Predictor significance and BCa confidence intervals were derived using balanced bootstrap (R=1000 resampling). Results were expressed as standardised regression coefficients  $\beta$  and credible intervals derived from the bias corrected accelerated bootstrap distribution. Coefficients with bootstrap distributions above or below zero were considered statistically significant.

## 2.5 Comparison with humans

The perception of distress encoded in infant bonobo, chimpanzee, and human cries by adult human listeners has been the subject of a previous study (Kelly et al., 2017). However, in order to compare with the results obtained here with crocodiles, we reanalyzed the data from that previous study normalizing human ratings between 0 and 1 and using the acoustic variables presented here. We conducted the same analyses as described above for crocodiles (principal component analysis and partial least squares logistic regression, taking *meanINThuman* instead of *meanINTcroc*).

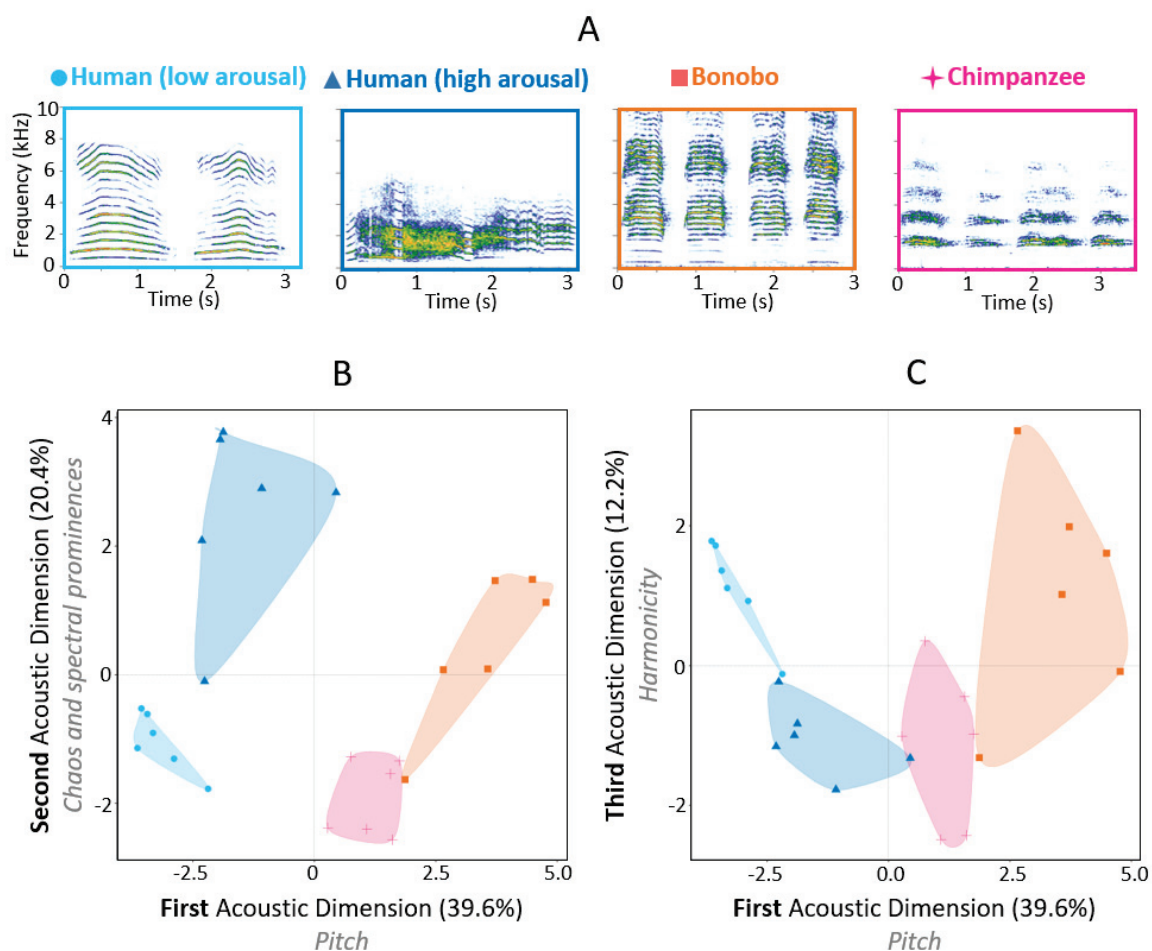
## 3. Results

### *Cry stimuli differ by pitch, presence of chaos and distribution of energy in the spectrum*

Principal Component Analysis performed on the 18 acoustic variables yields three major components (called “Acoustic Dimensions”) that significantly discriminate the cry categories (anova, AD1:  $F(3,20) = 70.6$ ,  $p < 0.001$ ; AD2:  $F(3,20) = 22.5$ ,  $p < 0.001$ ; AD3:  $F(3,20) = 8.4$ ,  $p < 0.001$ ). The first acoustic dimension (AD1) primarily represents cry pitch (with *meanF0*,



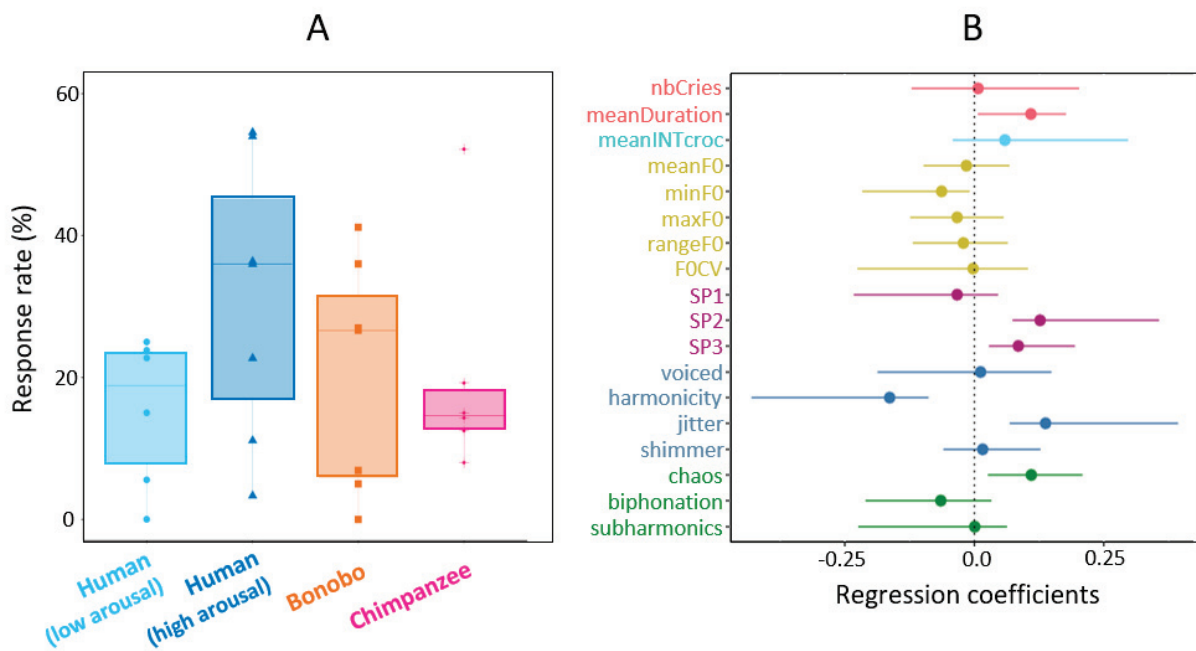
*maxF0*, and *rangeF0* showing the highest loadings; appendix Supplementary Table V.2). As illustrated by Figure V.1B, bonobo cries have a higher pitch than those of chimpanzees and human babies. Human babies recorded during bathing were the lowest pitched cries (see appendix Supplementary Table V.3 for post-hoc multiple comparisons between cry categories). The second acoustic dimension AD2 essentially represents deterministic chaos (nonlinear phenomenon) and the highest spectral prominences (SP2 and SP3). It is the cries of human babies recorded in a vaccination context and the cries of bonobos that present the highest values of AD2 (Figure V.1B). The third dimension (AD3) is associated with the periodic quality of the cries (captured by the variables *voiced*, *harm* and *jitter*; Figure V.1C).



**Figure V.1. Hominid baby cries used as playback stimuli.** A. Spectrograms of cry samples. B.C. Acoustic space of cries. Each dot represents a cry stimulus (duration around 3 seconds). The acoustic structure of cries was described using 18 acoustic variables and further reduced into three independent acoustic dimensions using a Principal Component Analysis. The first acoustic dimension (AD1) is mainly related to cry pitch. AD2 is mainly related to deterministic chaos and high spectral prominences. AD3 is mainly related to cry harmonicity (see appendix Supplementary Table V.2 for Principal Components coefficients).

### *Crocodile response is driven by a set of acoustic features that do not include pitch*

The results of the playback experiments suggest that crocodiles respond more or less intensely to different categories of cries. As shown in Figure V.2A, high arousal cries from human infants thus tend to induce a response more frequently than low arousal cries or those from chimpanzees. However, these differences in response are not statistically significant (GLM, Wald  $X^2 = 5.0$ ,  $p = 0.173$ ).



**Figure V.2: Behavioral reaction of adult crocodiles to the cries of hominid babies.** **A.** Crocodile response rates to sound stimuli. Playbacks were performed to groups of adult crocodiles from different ponds. Each group was tested with all stimuli in several successive experimental trials. Each dot in the panel represents an experimental trial. The outcome measured was the proportion of responding individuals in each tested crocodile group. **B.** Acoustic predictors of crocodile response to infant cries. Coefficients with bootstrap distributions above or below zero are statistically significant predictors. The crocodile response to sound stimuli is essentially predicted by harmonic, jitter, the presence of chaos and the highest spectral prominences. Predictors related to stimulus pitch do not explain the crocodile response (Partial Least Square Regression: standardized regression coefficients and 95% balanced bias-corrected and accelerated bootstrap confidence intervals).

The PLS regression reveals the acoustic predictors of crocodile response to sound stimuli (Figure V.2B). Low harmonic, high jitter, the presence of chaos, and higher energy in the



higher frequencies of the spectrum accompany higher responsiveness of the tested animals (harmonicity:  $\beta = -0.16$  [-0.43, -0.09]; jitter:  $\beta = 0.14$  [0.07, 0.39]; deterministic chaos:  $\beta = 0.11$  [0.03, 0.21]; SP2:  $\beta = 0.13$  [0.07, 0.36]; SP3:  $\beta = 0.09$  [0.03, 0.19]). Conversely, the pitch (F0) does not predict crocodile reaction to sound stimuli.

### ***Human listeners rely mostly on pitch features to assess baby cries***

Similar to the results reported in Kelly (2017), we found that human listeners judge the cry of hominid babies differentially (GLM, Wald  $X^2 = 92.0$ ,  $p < 0.001$ , appendix Supplementary Figure V.3A). Specifically, they rate bonobo cries as expressing the highest level of distress (multiple comparisons:  $\beta > 1.0$ ,  $Z > 3.5$ ,  $p < 0.01$ ), while human babies' low arousal cries are rated as expressing the least distress (multiple comparisons:  $\beta < -1.2$ ,  $Z < -5.4$ ,  $p < 0.001$ ).

PLS regression reveals that the most significant predictors are related to pitch and its variation (rangeF0:  $\beta = 0.08$  [0.07, 0.09]; maxF0:  $\beta = 0.07$  [0.06, 0.08]; F0CV:  $\beta = 0.08$  [0.06, 0.10]). Human listeners thus assign a high distress value to high-pitched cries. Other predictors, such as harmonicity ( $\beta = -0.07$  [-0.08, -0.06]), also modulate human listeners' rating of cries (appendix supplementary Figure V.3B).

## **4. Discussion**

Our results show that Nile crocodiles are attracted to infant hominid cries and suggest that the intensity of their response depends on acoustic features known to encode the intensity of distress expressed by the emitter. In particular, crocodiles are more attracted to cries with nonlinear acoustic phenomena (chaos, low harmonicity) and more intense energy in the high frequencies of the spectrum (spectral prominences). Unlike humans who primarily use pitch to judge the level of distress encoded in infant cries, crocodiles are only moderately sensitive to this acoustic feature.

The present analysis of the acoustic structure of bonobo, chimpanzee, and human infant cries confirms and complements the work of Kelly et al. (Kelly et al., 2017). The acoustic parameter differing the most between stimulus categories is pitch. Bonobo infant cries are by far the highest pitched, followed by human infant high arousal cries. Human baby high arousal cries are distinguished from other stimuli by a greater presence of nonlinear phenomena (Koutseff et al., 2018; Raine et al., 2019) (deterministic chaos, low harmonicity). As illustrated by the representation of the stimuli in two-dimensional acoustic space, some of the bonobo cries also

exhibit marked nonlinear phenomena (Waal, 1988) (high values of the second acoustic dimension). Because of their diversity, our stimuli thus represent a range of distress coding. Human infants were recorded in two contexts (at the bath and during a vaccination session) inducing different levels of arousal. Bonobo and chimpanzee infants were recorded in a variety of contexts, also inducing a diversity of arousal levels.

This diversity of arousal, which translates into a variation in the acoustic features of the stimuli, is accompanied by a variation in the response intensity of the crocodiles. Playback experiments show that crocodiles are not particularly sensitive to the category of the playback call (bonobo, chimpanzee or human), but pay particular attention to acoustic features that mark arousal, such as nonlinear acoustic phenomena (chaos and low harmonicity) and spectral prominences in the upper part of the spectrum. What might be termed the "perceptual dimension" of crocodiles does not correspond exactly to the distribution of calls in the acoustic space determined by principal component analysis. In particular, while it is an essential dimension for discriminating between cry categories, pitch is not a reliable predictor of crocodile response.

This result is interesting for two reasons. First, it marks a difference with the way humans assess the level of distress in infant cries. Our analysis of humans' assignment of distress level to stimuli indeed confirms that conducted by Kelly et al. (Kelly et al., 2017): humans assign a distress level primarily from the pitch of the cry. The higher the pitch of a cry, the more humans judge the cry as expressing high distress (McComb et al., 2009; Faragó et al., 2014; Maruščáková et al., 2015; Filippi et al., 2017). This leads human listeners to consistently judge bonobo baby cries as expressing strong distress, and thus to be likely to be wrong. Bonobo babies have been recorded in a variety of contexts and express a diversity of arousal levels, which is reflected in the wide distribution of their cries on the second acoustic dimension (chaos and spectral prominences). The second interest of the fact that crocodiles do not pay attention to pitch is that their reaction to the cry of babies is less dependent on the species of animal emitting the cry than for humans. Pitch is a potentially misleading distress marker trait: while it may be informative within a given animal species, its basal value varies too much between animal species to be a universal marker of distress (Bremond, 1976; Root-Gutteridge et al., 2021; Kelly et al., 2017; Teichroeb et al., 2013; Lingle & Riede, 2014). Crocodiles thus seem particularly adapted at estimating the degree of distress encoded in an infant's cry regardless of the hominid species considered. Unlike humans, whose perception and analysis of infant cries is biased by this emphasis on pitch, crocodiles probably have no experience with cries in different hominid species - except in the wild for crocodiles living in close proximity to human





populations or other Hominids. They therefore respond to stimuli based on acoustic criteria alone, without recognizing the origin of the stimulus. Moreover, although crocodiles have excellent hearing in air, they are mostly sensitive between 100 and 2000-3000 Hz (Higgs et al., 2002; A. Vergne et al., 2009). Bonobo cries, which have a frequency range up to 10 kHz, are therefore likely to be poorly perceived by crocodiles. This poor perception of high frequencies may explain why human babies' high arousal cries tend to induce a stronger response from crocodiles.

Why humans attribute so much importance to the pitch of cries in judging the level of distress is a question beyond the scope of this paper. To put it in a nutshell, this is potentially explained by the fact that a human infant expresses a high level of distress by increasing the prevalence of not only nonlinear acoustic phenomena but also pitch (Koutseff et al., 2018). The average pitch differs between human babies, so pitch is a poor universal marker of distress, but since a human normally knows the baby(s) they are caring for, using pitch to detect distress in a given baby's cries becomes reliable.

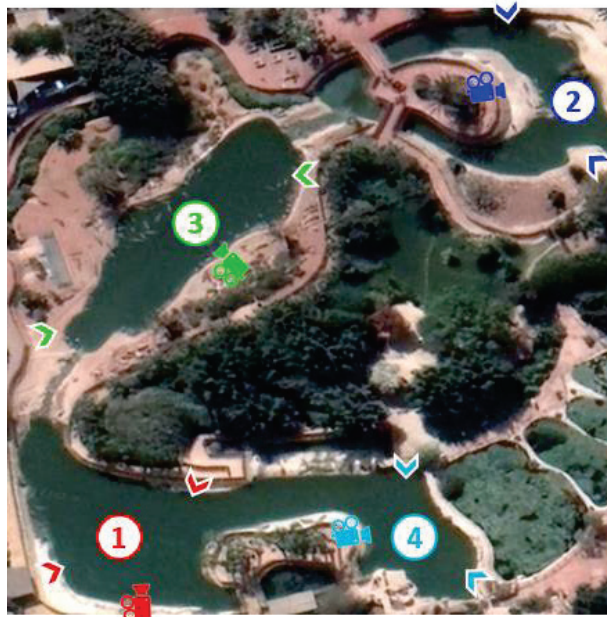
Why do crocodiles respond to hominid cries? It is known that adult crocodiles are attracted to the distress calls of their young. Crocodylian females -and males in some species- come to provide parental assistance in case of attack by a predator (Somaweera et al., 2013). In other species, males practice cannibalism (Hutton, 1989; Grigg & Kirshner, 2015). The distress calls of young crocodylians share common features with the cries of hominid infants: they are harmonic series spanning a wide bandwidth (between 500 and 6000 Hz) (A. Vergne et al., 2007; A. L. Vergne et al., 2011), modulated in frequency, and may exhibit nonlinear acoustic phenomena. In our experiments, it was not possible to reliably identify each tested adult crocodile as male or female. However, we definitely observed that individuals of both sexes responded to our stimuli. In addition, not all animals approached the speaker in the same way. Some did so by swimming on the surface, while others practiced an underwater approach then tried to bite the speaker, a behavior which looks as a predator strategy. We also observed some individuals stopping a few tens of centimeters in front of it with their snout pointed in its direction. It is therefore likely that our hominid cry stimuli triggered both types of responses from the crocodiles, parental care and predation, depending on the individual. In any case, of the two behaviors, it is far more likely that the crocodiles perceive distress calls as indicating easy prey. This might suggest that they would respond to the distress calls of other species because they might see this as a possible food source. Because of their population ecology - their proximity to human/non-human primate groups in some parts of the world now and in the

past, crocodylians' sensitivity to prey distress would be of adaptive interest and used to hunt our early ancestors. Further, we can hypothesize that crocodylians are additionally able to perceive the level of distress and thus direct their attention to the easiest prey representing the best hunting opportunities.

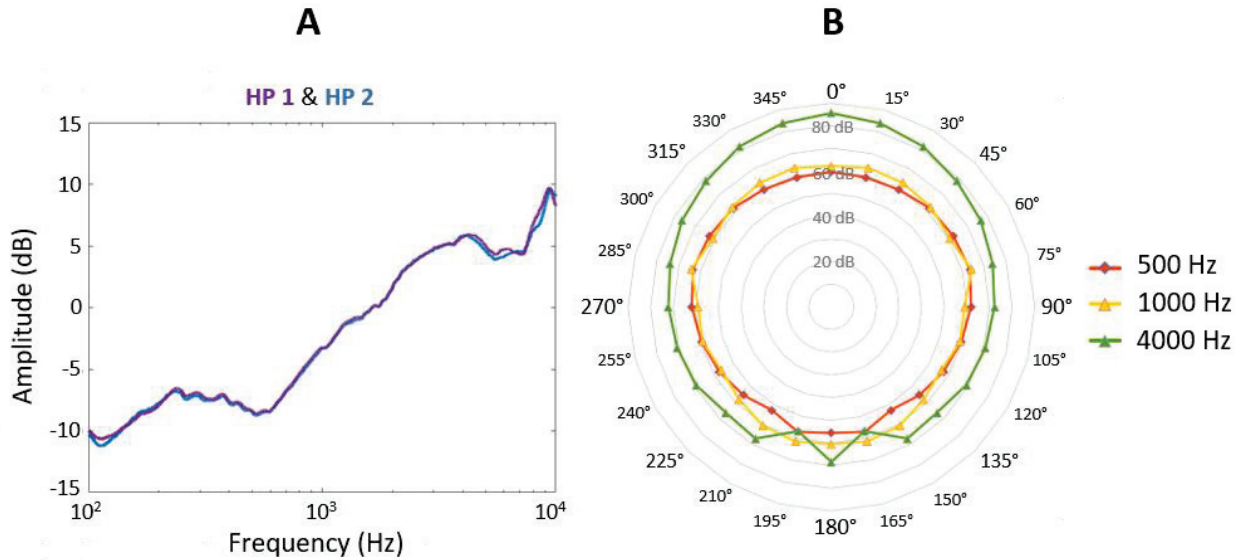
In sum, our study shows that crocodiles, for perhaps complementary reasons (parental care versus predation) identify and respond proportionately to different levels of distress in hominid infant cries. This behavior is elicited by acoustic features otherwise known to be markers of distress in vocalizations (Lingle et al., 2012; Massenot et al., 2022; Anikin et al., 2021, 2020; Marx et al., 2021; Rendall et al., 2009; Stoeger et al., 2011). Because the crocodiles tested have never had the opportunity to associate a hominid baby's cry with the presence of a baby in their immediate surroundings in the zoo, their response to hominid baby cries is likely innate. These results suggest that the supposed universality of acoustic features encoding distress in vocalizations does exist in the vertebrate group, and that these features are primarily represented by nonlinear acoustic phenomena rather than by the pitch of vocalizations.

## 5. Supplementary information

**Supplementary Material V.1. Conversion of flat dB to dB “crocodile”.** Based on the ANSI methodology used to estimate the sound level in dB(A) (i.e. related to the human auditory curve; ANSI S1.6-1967 (R1976)) (American National Standards Institute & United States of America Standards Institute, 1967), we developed the same methodology based on the alligator audiogram as measured by Higgs et al. (2002) to estimate the sound intensity in dB “crocodile”. The first step was to interpolate a weighting function in dB for each frequency from the alligator audiogram curve (ref 0 dB at 1000Hz) and to transform this weighting function on a linear scale. The second step was to compute the spectrum of the sounds in dB using an FFT (Fast Fourier Transform) and then to multiply this FFT transformed on linear scale by the weighting function. Finally, we calculated the mean RMS level in dB of the resulting FFT.



**Supplementary Figure V.1. Aerial view of the ponds of Crocoparc Agadir, Morocco.** We played the sound stimuli to four groups of 7 to 25 adults (females and males, sex ratio unknown), occupying 4 different ponds. For each experimental session, two loudspeakers were placed on the banks of the pond, allowing to playback the sound stimuli from two different locations. Arrows indicate the position and direction of the loudspeakers. Camera icons indicate the position of the camera.



**Supplementary Figure V.2. Technical specifications of the loudspeakers.** Most of the energy of the acoustic signals used in this study was between 500 Hz and 10 kHz. Since the original transfer functions of the FoxPro loudspeakers (Shogyo #GF0923BM-1X) in this frequency range were very hilly with rapid variations up to 16 dB, we decided to insert a new speaker into the FoxPro with smoother and flatter transfer functions (Visaton #SL 87 ND). In addition, the use of these new loudspeakers makes the FoxPro much more omnidirectional. All measurements have been done in a semi-anechoic room (dimensions 3.4 x 4.6 m; reverberation time = 0.2 for 125 Hz and  $\leq 0.1$  for frequencies greater than or equal to 500 Hz). **A.** Transfer functions of the two FoxPro (1 and 2) with the new loudspeakers measured with a condenser microphone (Behringer ECM8000) while emitting a broadband noise. **B.** Directionality functions (polar diagram, dB scale) of the FoxPro with the new internal speakers measured for sinus at three frequencies.

<b>Pond</b>	<b>Trial</b>	<b>Cry category</b>	<b>Signal ID</b>	<b>Nb of individuals in the pond</b>	<b>Response rate</b>
1	1	bonobo	03_bonobo	25	36 %
1	2	chimpanzee	04_chimp	20	15 %
1	3	Human bath	01_human_bath	18	6 %
1	4	bonobo	03_bonobo	22	27 %
1	5	Human vaccine	02_human_vaccine	11	55 %
1	6	Human vaccine	02_human_vaccine	28	54 %
2	1	Human bath	06_human_bath	7	0 %
2	2	Human vaccine	11_human_vaccine	11	36 %
2	3	bonobo	16_bonobo	12	0 %
2	4	chimpanzee	21_chimp	8	13 %
3	1	chimpanzee	22_chimp	23	52 %
3	2	Human vaccine	12_human_vaccine	30	3 %
3	3	bonobo	17_bonobo	30	27 %
3	4	Human bath	07_human_bath	21	24 %
3	5	bonobo	18_bonobo	29	7 %
3	6	chimpanzee	23_chimp	21	14 %
4	1	Human bath	08_human_bath	20	25 %
4	2	bonobo	19_bonobo	17	41 %
4	3	Human vaccine	13_human_vaccine	22	23 %
4	4	chimpanzee	24_chimp	25	8 %
4	5	Human bath	09_human_bath	20	15 %
4	6	Human vaccine	14_human_vaccine	18	11 %
1	1	Human vaccine	15_human_vaccine	25	36 %
1	2	chimpanzee	25_chimp	26	19 %
1	3	bonobo	20_bonobo	20	5 %
1	4	Human bath	10_human_bath	22	23 %

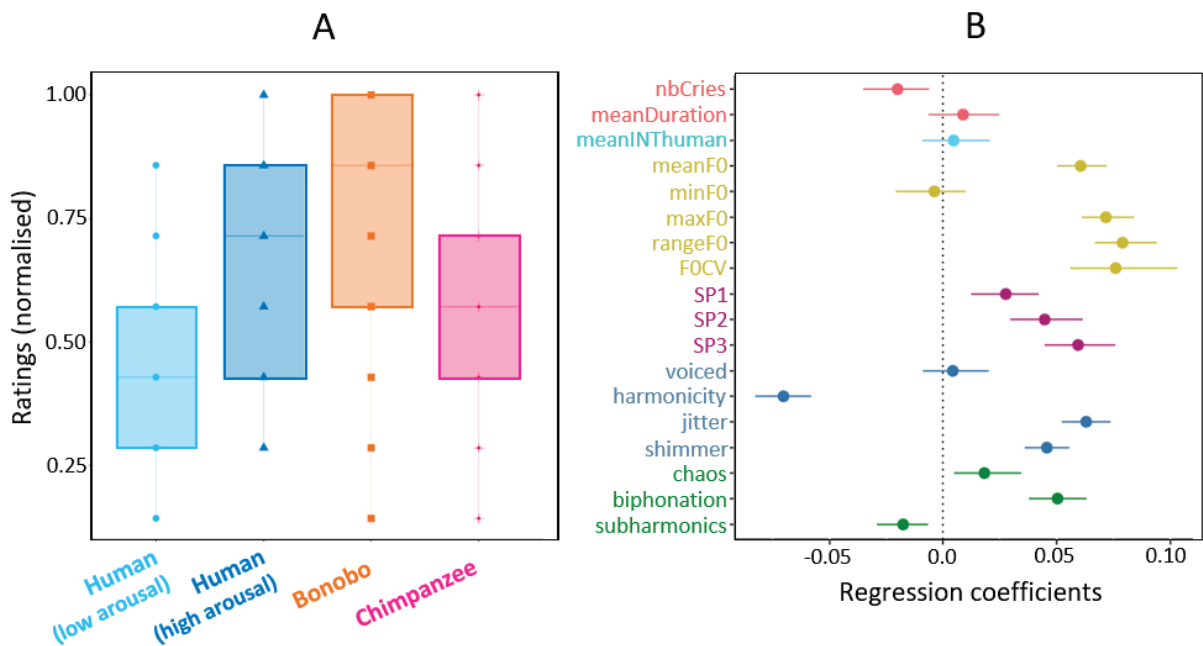
**Supplementary Table V.1. Details of the playback experiments.**

<b>Acoustic parameters</b>	<b>AD1 Variance = 39.6% Eigenvalue = 7.1</b>	<b>AD2 Variance = 20.4% Eigenvalue = 3.7</b>	<b>AD3 Variance = 12.2% Eigenvalue = 2.2</b>
<i>nbCries</i>	0.35	-0.65	-0.33
<i>meanDuration</i>	-0.44	0.69	-0.07
<i>voiced</i>	0.08	-0.05	<b><u>0.72</u></b>
<i>meanFO</i>	<b><u>0.91</u></b>	-0.17	0.05
<i>maxFO</i>	<b><u>0.94</u></b>	-0.16	0.18
<i>minFO</i>	0.31	-0.51	-0.31
<i>rangeFO</i>	<b><u>0.91</u></b>	-0.04	0.27
<i>FOCV</i>	0.42	0.44	0.48
<i>meanINTcroc</i>	-0.68	-0.41	-0.32
<i>harmonicity</i>	-0.62	-0.27	<b><u>0.54</u></b>
<i>jitter</i>	0.77	0.08	<b><u>-0.54</u></b>
<i>shimmer</i>	0.79	-0.21	-0.43
<i>SP1</i>	0.70	0.41	-0.03
<i>SP2</i>	0.47	<b><u>0.75</u></b>	-0.03
<i>SP3</i>	0.48	<b><u>0.74</u></b>	-0.12
<i>biphonation</i>	0.84	-0.14	0.32
<i>subharmonics</i>	-0.37	0.22	-0.03
<i>chaos</i>	-0.35	<b><u>0.79</u></b>	-0.40

**Supplementary Table V.2. Principal Component Analysis of the sound stimuli.**

<i>Cry categories comparison</i>	<i>Acoustic Dimension 1</i>			<i>Acoustic Dimension 2</i>			<i>Acoustic Dimension 3</i>		
	Estimate	t	p	Estimate	t	p	Estimate	t	p
<i>Human bath/ Human vaccine</i>	-1.66	-3.36	<b>0.015</b>	-3.57	-6.16	<b>&lt;0.001</b>	2.18	3.49	<b>0.011</b>
<i>Human bath/ bonobo</i>	-6.68	-13.48	<b>&lt;0.001</b>	-1.48	-2.55	0.081	0.03	0.05	1.000
<i>Human bath/ chimpanzee</i>	-4.34	-8.76	<b>&lt;0.001</b>	0.87	1.50	0.454	2.29	3.67	<b>&lt;0.01</b>
<i>Human vaccine / bonobo</i>	-5.02	-10.12	<b>&lt;0.001</b>	2.09	3.61	<b>&lt;0.01</b>	-2.15	-3.44	<b>0.013</b>
<i>Human vaccine / chimpanzee</i>	-2.68	-5.40	<b>&lt;0.001</b>	4.44	7.66	<b>&lt;0.001</b>	0.11	0.18	1.000
<i>Bonobo/ chimpanzee</i>	2.34	4.72	<b>&lt;0.001</b>	2.35	4.06	<b>&lt;0.01</b>	2.26	3.62	<b>&lt;0.01</b>

**Supplementary Table V.3. Multiple comparisons tests between sound stimuli.**



**Supplementary Figure V.3. Ratings of distress level in hominid babies by human adults.**

**A.** Normalised distress level ratings. Each dot represents one psychoacoustic measure. **B.** Acoustic predictors of human rating of distress level in infant cries. Coefficients with bootstrap distributions above or below zero are statistically significant predictors. Predictors related to stimulus pitch and its variation explain the best human ratings (Partial Least Square Regression: standardised regression coefficients and 95% balanced bias-corrected and accelerated bootstrap confidence intervals).

## VI. General discussion

The goal of my PhD works was to better understand how crocodylians perceive their sound environment and how they use acoustic information to make decisions about whether to respond to a signal. Since the sound environment in which crocodylians live is complex, even noisy, I was first interested in the ability of these animals to detect a sound signal masked by noise. Then, once this signal is detected, I tried to understand how crocodylians analyze a set of sound signals and if they are able to partition the information that they receive into meaningful categories. Finally, I studied the ability of crocodylians to use acoustic information that is not directly intended for them but that might be of interest for them, focusing on their perception of distress encoded in vocalizations of distant species.

When we are confronted with a noise masking the sound signal of interest, spatial unmasking is one of the mechanisms that can allow us to better detect this signal and that has been demonstrated in several species (especially birds). We therefore studied this phenomenon through a multi-contextual approach. We first studied this phenomenon in the wild, by observing the response of female Yacare caimans to distress calls masked by white noise. The results were more than conclusive, showing that females lowered their response threshold when distress calls were emitted from a source distinct from the noise. Next, spatial unmasking was explored in young Nile crocodiles in a zoo using contact calls (attraction function). Again, juveniles responded more readily to calls that were made far from the white noise source. Finally, two Nile crocodiles were trained in the laboratory to respond to synthetic signals. Their threshold for detecting these same signals appeared to be lowered when they were emitted spatially separate from the noise source, although the results were not as clear.

While we would have thought that laboratory experiments would refine our observations, instead, multiple constraints complicated the study of this phenomenon. The crocodiles in the laboratory were not at risk of not responding to a signal they had detected as might be the case for animals in natural conditions. It was therefore difficult in this experiment to maintain a sufficiently constant motivation of the individuals in order to obtain stable results. Moreover, the experimental set-up in the laboratory also brought its share of constraints, notably the fact that we are in a closed environment, thus subject to strong reverberation. The fact that animals are at risk of not responding to a signal at the end of the day in a natural environment can also work in the other direction: they can also risk responding to a signal that they have just detected





and potentially misidentified. Crocodylians may face numerous trade-off situations that have an impact on the responses that have been measured (yacara females must decide between responding and helping a potential pup in distress in the distance or staying with the group of pups around her; young Nile crocodiles between responding and moving to a potential conspecific or staying put and remaining as safe as possible) In these studies, therefore, one would refer more to a single behavioral response to a signal rather than to absolute thresholds of signal detection. Inter-individual differences can explain the fluctuation of the results obtained. These differences can be first in terms of species between the different experiments. Nile crocodiles are known to be more reckless and much more aggressive than yacare caimans. Therefore, they would need less confidence before deciding to respond to a sound signal. Personality differences can also exist within the same species. Bolder and more exploratory individuals will also tend to lower their behavioral response threshold.

However, it is not enough for the signal to be detected for the animals to respond. A threshold of sound intensity is certainly necessary to be crossed in order for the signal to be detectable, but it is also necessary for this signal to contain enough information for it to be identified as being of interest and classified as having this or that meaning. To explore this informational threshold, I conducted categorization experiments using an artificial acoustic continuum between a crocodylian contact call and a frog call. A preliminary experiment was conducted on young Nile crocodiles in a zoo. The results show that the juveniles need a significant amount of acoustic information in the call (almost all of it) to be able to respond. The experiments were then taken further in the laboratory by training young dwarf caiman in a go/no-go paradigm: they were trained to come under a speaker emitting a crocodile call and to stay away from it when it emitted a frog call. This allowed me to test a range of chimerical signals between the crocodile and frog call, to see to which category they attribute it and whether they discriminate better between calls belonging to different perceptual categories than calls from the same category (the two criteria for attesting to a categorization process). I show that caimans do categorize the acoustic continuum into two distinct categories (crocodile category and frog category) based almost solely on the spectral envelope of calls. Compared to results obtained with naive young crocodiles, the boundary between the two categories is largely shifted to a lower information threshold. These results are consistent with the same conclusion drawn from the zoo spatial unmasking experiment where juveniles are placed in such an isolated and potentially stressful situation that they likely need to raise their behavioral response thresholds in order to be sure they are not mistaken. Once again, the thresholds measured will be strongly

dependent on the context in which the animal is, but also on its age, its experience, and even its personality.

To lower the cost to make errors, animals will respond only if the signal is within a defined category with a sufficient probability (Akre & Johnsen, 2014; Bee & Miller, 2016). Error in classification and responses can also be due to *small just-noticeable difference* (JND, the minimum difference between two signals that can be detected): the larger the proportional differences between two stimuli is, the better their discrimination is. This was observed when using signals with a smaller information step (continuum of signals every 5%) for which the animals were less constant in their classification. Nevertheless, overall in the laboratory, the thresholds remained constant and the same over time and for all individuals tested. This can be explained by the fact that the animals were trained to the task and thus had no trade-offs to overcome. However, differences in individuality were expressed during the conditioning process, which necessitated a progressive and individualized evolution of the conditioning steps.

The categorization experiments were complemented by pupillometry experiments to get a more accurate indication of signal discrimination performance. Based on the oddball effect principle where a signal is repeated until habituation, we hypothesized that the pupil would dilate in response to the sudden presentation of a new signal (orienting reflex) and that signals belonging to different perceptual categories would elicit a greater response from the pupil because they would be better discriminated than those belonging to the same category. However, the implementation of a system adapted to the tracking of the caiman pupil, including the fact that they had to be placed under restraint, was long and difficult. These difficulties probably tarnish the results: they do not allow us to conclude that there is a pupillary response to the oddball effect in crocodylians, nor to confirm our hypotheses. However, the principle remains promising: the time spent with these animals and the conditioning sessions allowed us to observe that the caimans have pupils that dilate rapidly under the effect of a stimulation such as the sudden emission of a sound signal of interest (like the signals used for the conditioning to which they associated the food). Interestingly, the individuals who were the least social and most aggressive in general over the three years were those who also showed the greatest pupillary dilation when stimulated in conditioning. Individuals with calmer traits would therefore potentially either be less motivated by food or have reduced norepinephric activity, which would not be surprising when one assumes that catecholamine deliberation affects behavior and especially aggression (Bell & Hepper, 1987; Haller et al., 1998). The methodology



around this project would therefore deserve to be further developed in order to have an optimal tracking system for the crocodile pupil.

If we go back to the story, crocodilians are able to detect sound efficiently even though their environment may be noisy. They also have perceptual and cognitive mechanisms that allow them to identify these sounds and classify them into functional categories. The last question was therefore to know if they are also able to take advantage of sounds that are not a priori intended for them, by taking as a subject of study their perception of distress calls produced by other animals. We show through playback experiments that crocodiles are attracted to distress calls produced by different hominid species and that their response is modulated by the acoustic characteristics of the calls. The roughest calls elicited the strongest response from crocodiles, suggesting that they are quite good at perceiving the level of distress encoded in a call since roughness parameters (such as the presence of deterministic chaos in the signal) have often been associated with a high level of arousal. These results highlight that crocodilians are able to decode heterospecific information, using acoustic parameters that seem to be universal markers of distress, but also that crocodilians can potentially use this information in their environment to optimize their hunting strategy and assert their top predator status. However, these results remain preliminary and would require further experiments. The creation of synthetic signals would allow us to account for the real effect of each acoustic parameter independently of the others, by modifying only the parameters of interest, keeping all the other parameters identical.

In sum, we can conclude from this work that crocodilians have a fine perception of their sound environment and that they are able to efficiently extract and identify sound signals in an environment that is nevertheless rich in information. In all the experiments I conducted during my thesis, the environmental factor played a major role in the results obtained. Working on animal perception is not as obvious as working on human perception. Even if an animal is able to detect and then identify and classify a signal as a signal of interest, it faces multiple external constraints and must take into account all these factors before making a decision. In the laboratory where animals are trained to respond to a particular task, stress, satiety and many other physiological parameters can also impact their behavior and performance. In the wild, many trade-offs can modulate the choice of an animal, especially when a young predator is involved. If we have seen that these animals are able to categorize signals along an acoustic continuum, it would now be interesting to study in more detail whether this mechanism of categorical perception also helps crocodilians to make decisions in other contexts and applies to the perception of calls from their own vocal repertoire. For example, juvenile contact and

distress calls have very distinct (even opposite) functions and effects, yet their acoustic structure is gradual from one call to the next. Thus, it could be hypothesized that both juvenile crocodilians and mothers categorically perceive this continuum and rely on a very specific amount of acoustic information to classify calls as either contact or distress calls.

Finally, more generally, these studies have opened the door to new directions of research on crocodilians. The questions of learning, memory, or personality traits that have arisen throughout my experiments are all subjects that we can dig up in crocodilians and that will certainly bring once again their share of surprises.





# References

- Akre, K. L., & Johnsen, S. (2014). Psychophysics and the evolution of behavior. *Trends in Ecology & Evolution*, 29(5), 291–300. <https://doi.org/10.1016/j.tree.2014.03.007>
- Akre, K. L., & Johnsen, S. (2016). Communication Through a Window of Error: Proportional Processing and Signal Categorization. In M. A. Bee & C. T. Miller (Eds.), *Psychological Mechanisms in Animal Communication* (pp. 137–167). Springer International Publishing. [https://doi.org/10.1007/978-3-319-48690-1\\_6](https://doi.org/10.1007/978-3-319-48690-1_6)
- Alamia, A., VanRullen, R., Pasqualotto, E., Mouraux, A., & Zenon, A. (2019). Pupil-Linked Arousal Responds to Unconscious Surprisal. *Journal of Neuroscience*, 39(27), 5369–5376. <https://doi.org/10.1523/JNEUROSCI.3010-18.2019>
- American National Standards Institute & United States of America Standards Institute. (1967). *Preferred frequencies and band numbers for acoustical measurements*. United States of America Standards Institute.
- Anikin, A., Pisanski, K., Massenet, M., & Reby, D. (2021). Harsh is large: Nonlinear vocal phenomena lower voice pitch and exaggerate body size. *Proceedings of the Royal Society B: Biological Sciences*, 288(1954), 20210872. <https://doi.org/10.1098/rspb.2021.0872>
- Anikin, A., Pisanski, K., & Reby, D. (2020). Do nonlinear vocal phenomena signal negative valence or high emotion intensity? *Royal Society Open Science*, 7(12), 201306. <https://doi.org/10.1098/rsos.201306>
- Aubin, T. (1991). Why do distress calls evoke interspecific responses? An experimental study applied to some species of birds. *Behavioural Processes*, 23(2), 103–111. [https://doi.org/10.1016/0376-6357\(91\)90061-4](https://doi.org/10.1016/0376-6357(91)90061-4)
- Aubin, T., & Jouventin, P. (2002). Localisation of an acoustic signal in a noisy environment: The display call of the king penguin *Aptenodytes patagonicus*. *The Journal of Experimental Biology*, 205, 3793–3798. <https://doi.org/10.1242/jeb.205.24.3793>
- August, P. V., & Anderson, J. G. T. (1987). Mammal Sounds and Motivation-Structural Rules: A Test of the Hypothesis. *Journal of Mammalogy*, 68(1), 1–9. <https://doi.org/10.2307/1381039>
- Baird, I. L. (1974). Anatomical Features of the Inner Ear in Submammalian Vertebrates. In H. W. Ades, A. Axelsson, I. L. Baird, G. v. Békésy, R. L. Boord, C. B. G. Campbell, O. Densert, D. H. Eldredge, H. Engström, J. Fex, J. M. Harrison, O. W. Henson, M. E. Howe, S. Iurato, A. Michelsen, A. R. Møller, R. R. Pfeiffer, S. Rauch, I. Rauch, ... W. D. Neff (Eds.), *Auditory System: Anatomy Physiology (Ear)* (pp. 159–212). Springer. [https://doi.org/10.1007/978-3-642-65829-7\\_6](https://doi.org/10.1007/978-3-642-65829-7_6)
- Bala, A. D. S., & Takahashi, T. T. (2000). Pupillary dilation response as an indicator of auditory discrimination in the barn owl. *Journal of Comparative Physiology A*, 186(5), 425–434. <https://doi.org/10.1007/s003590050442>
- Barbur, J. L., Prescott, N. B., Douglas, R. H., Jarvis, J. R., & Wathes, C. M. (2002). A comparative study of stimulus-specific pupil responses in the domestic fowl (*Gallus gallus domesticus*) and the human. *Vision Research*, 42(2), 249–255. [https://doi.org/10.1016/S0042-6989\(01\)00279-6](https://doi.org/10.1016/S0042-6989(01)00279-6)



- Bard, K. A. (2000). Crying in infant primates: Insights into the development of crying in chimpanzees. In *Crying as a sign, a symptom, & a signal: Clinical emotional and developmental aspects of infant and toddler crying* (pp. 157–175). Cambridge University Press.
- Bee, M. A. (2007). Sound source segregation in grey treefrogs: Spatial release from masking by the sound of a chorus. *Animal Behaviour*, *74*(3), 549–558. <https://doi.org/10.1016/j.anbehav.2006.12.012>
- Bee, M. A. (2008). Finding a mate at a cocktail party: Spatial release from masking improves acoustic mate recognition in grey treefrogs. *Animal Behaviour*, *75*(5), 1781–1791. <https://doi.org/10.1016/j.anbehav.2007.10.032>
- Bee, M. A., & Micheyl, C. (2008). The cocktail party problem: What is it? How can it be solved? And why should animal behaviorists study it? *Journal of Comparative Psychology*, *122*(3), 235–251. <https://doi.org/10.1037/0735-7036.122.3.235>
- Bee, M. A., & Miller, C. T. (2016). Signaler and Receiver Psychology. In M. A. Bee & C. T. Miller (Eds.), *Psychological Mechanisms in Animal Communication* (pp. 1–16). Springer International Publishing. [https://doi.org/10.1007/978-3-319-48690-1\\_1](https://doi.org/10.1007/978-3-319-48690-1_1)
- Bee, M. A., & Schwartz, J. J. (2009). Behavioral measures of signal recognition thresholds in frogs in the presence and absence of chorus-shaped noise. *The Journal of the Acoustical Society of America*, *126*(5), 2788–2801. <https://doi.org/10.1121/1.3224707>
- Bee, M. A., & Vélez, A. (2018). Masking release in temporally fluctuating noise depends on comodulation and overall level in Cope's gray treefrog. *The Journal of the Acoustical Society of America*, *144*(4), 2354–2362. <https://doi.org/10.1121/1.5064362>
- Bermejo, M., & Omedes, A. (2000). Preliminary Vocal Repertoire and Vocal Communication of Wild Bonobos (*Pan paniscus*) at Lilungu (Democratic Republic of Congo). *Folia Primatologica*, *70*(6), 328–357. <https://doi.org/10.1159/000021717>
- Bertrand, F., Maumy, M., & Meyer, N. (2009). PlsRglm, modèles linéaires généralisés PLS sous R. *Chimiométrie 2009*, pp 52-54. <https://hal.archives-ouvertes.fr/hal-00439913>
- Bierman, H. S., & Carr, C. E. (2015). Sound localization in the alligator. *Hearing Research*, *329*, 11–20. <https://doi.org/10.1016/j.heares.2015.05.009>
- Bierman, H. S., Thornton, J. L., Jones, H. G., Koka, K., Young, B. A., Brandt, C., Christensen-Dalsgaard, J., Carr, C. E., & Tollin, D. J. (2014). Biophysics of directional hearing in the American alligator (*Alligator mississippiensis*). *Journal of Experimental Biology*, *217*(7), 1094–1107. <https://doi.org/10.1242/jeb.092866>
- Blauert, J. (1997). *Spatial hearing: The psychophysics of human sound localization*. MIT press. <https://doi.org/10.7551/mitpress/6391.001.0001>
- Blumstein, D. T., Richardson, D. T., Cooley, L., Winternitz, J., & Daniel, J. C. (2008). The structure, meaning and function of yellow-bellied marmot pup screams. *Animal Behaviour*, *76*(3), 1055–1064. <https://doi.org/10.1016/j.anbehav.2008.06.002>
- Boersma, P., & Weenink, D. (1992). *PRAAT: Doing phonetics by computer (Version 6.2.06)*. <https://www.praat.org>
- Boucher, M., Tellez, M., & Anderson, J. T. (2020). Differences in distress: Variance and production of American Crocodile (*Crocodylus acutus*) distress calls in Belize. *Ecology and Evolution*, *10*(18), 9624–9634. <https://doi.org/10.1002/ece3.6556>

- Bradbury, J. W., & Vehrencamp, S. L. (2011). *Principles of animal communication, 2nd ed* (pp. xiv, 697). Sinauer Associates.
- Bradley, M. M., Miccoli, L., Escrig, M. A., & Lang, P. J. (2008). The pupil as a measure of emotional arousal and autonomic activation. *Psychophysiology*, *45*(4), 602–607. <https://doi.org/10.1111/j.1469-8986.2008.00654.x>
- Bremond, J.-C. (1976). Specific recognition in the song of Bonelli's warbler ( *Phylloscopus bonelli* ). *Behaviour*, *58*(1–2), 99–116. <https://doi.org/10.1163/156853976X00253>
- Briefer, E. F. (2012). Vocal expression of emotions in mammals: Mechanisms of production and evidence. *Journal of Zoology*, *288*(1), 1–20. <https://doi.org/10.1111/j.1469-7998.2012.00920.x>
- Briefer, E. F. (2020). Coding for 'Dynamic' Information: Vocal Expression of Emotional Arousal and Valence in Non-human Animals. In T. Aubin & N. Mathevon (Eds.), *Coding Strategies in Vertebrate Acoustic Communication* (pp. 137–162). Springer International Publishing. [https://doi.org/10.1007/978-3-030-39200-0\\_6](https://doi.org/10.1007/978-3-030-39200-0_6)
- Brien, M. L., Lang, J. W., Webb, G. J., Stevenson, C., & Christian, K. A. (2013). The Good, the Bad, and the Ugly: Agonistic Behaviour in Juvenile Crocodilians. *PLOS ONE*, *8*(12), e80872. <https://doi.org/10.1371/journal.pone.0080872>
- Brink, R. L. van den, Murphy, P. R., & Nieuwenhuis, S. (2016). Pupil Diameter Tracks Lapses of Attention. *PLOS ONE*, *11*(10), e0165274. <https://doi.org/10.1371/journal.pone.0165274>
- Britton, A. (2001). Review and classification of call types of juvenile crocodilians and factors affecting distress calls. *Crocodilian Biology and Evolution*, 364–377.
- Bronkhorst, A. W. (2000). The cocktail party phenomenon: A review of research on speech intelligibility in multiple-talker conditions. *Acta Acustica United with Acustica*, *86*(1), 117–128. <https://doi.org/10.3758/s13414-015-0882-9>
- Brumm, H., & Slabbekoorn, H. (2005). Acoustic communication in noise. *Advances in the Study of Behavior*, *35*, 151–209. [https://doi.org/10.1016/S0065-3454\(05\)35004-2](https://doi.org/10.1016/S0065-3454(05)35004-2)
- Burgering, M. A., Vroomen, J., & ten Cate, C. (2019). Zebra finches (*Taeniopygia guttata*) can categorize vowel-like sounds on both the fundamental frequency (“pitch”) and spectral envelope. *Journal of Comparative Psychology*, *133*(1), 106–117. <https://doi.org/10.1037/com0000143>
- Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, *80*, 1–28. <https://doi.org/10.18637/jss.v080.i01>
- Bürkner, P.-C., & Vuorre, M. (2018). Ordinal regression models in psychological research: A tutorial. *PsyArXiv Preprints*.
- Caldwell, M. S., Lee, N., & Bee, M. A. (2016). Inherent Directionality Determines Spatial Release from Masking at the Tympanum in a Vertebrate with Internally Coupled Ears. *Journal of the Association for Research in Otolaryngology*, *17*(4), 259–270. <https://doi.org/10.1007/s10162-016-0568-6>
- Campbell, H. D. (1973). Observations on the acoustic behavior of crocodilians. *Zoologica : Scientific Contributions of the New York Zoological Society*. <https://doi.org/10.5962/p.203220>





- Campos, Z., & Magnusson, W. (1995). Relationships between rainfall, nesting habitat and fecundity of *Caiman crocodilus yacare* in the Pantanal, Brazil. *Journal of Tropical Ecology*, *11*(3), 351–358. <https://doi.org/10.1017/S0266467400008828>
- Campos, Z., Mourao, G., Coutinho, M., Magnusson, W. E., & Soriano, B. M. A. (2015). Spatial and Temporal Variation in Reproduction of a Generalist Crocodylian, *Caiman crocodilus yacare*, in a Seasonally Flooded Wetland. *PLOS ONE*, *10*(6), e0129368. <https://doi.org/10.1371/journal.pone.0129368>
- Campos, Z., Sanaiotti, T., Muniz, F., Farias, I., & Magnusson, W. E. (2012). Parental care in the dwarf caiman, *Paleosuchus palpebrosus* Cuvier, 1807 (Reptilia: Crocodylia: Alligatoridae). *Journal of Natural History*, *46*(47–48), 2979–2984. <https://doi.org/10.1080/00222933.2012.724723>
- Carr, C. E., & Christensen-Dalsgaard, J. (2015). Sound Localization Strategies in Three Predators. *Brain, Behavior and Evolution*, *86*(1), 17–27. <https://doi.org/10.1159/000435946>
- Carr, C. E., Christensen-Dalsgaard, J., & Bierman, H. (2016). Coupled ears in lizards and crocodylians. *Biological Cybernetics*, *110*(4–5), 291–302. <https://doi.org/10.1007/s00422-016-0698-2>
- Carr, C. E., Soares, D., Smolders, J., & Simon, J. Z. (2009). Detection of Interaural Time Differences in the Alligator. *Journal of Neuroscience*, *29*(25), 7978–7990. <https://doi.org/10.1523/JNEUROSCI.6154-08.2009>
- Cavanagh, J. F., Wiecki, T. V., Kochar, A., & Frank, M. J. (2014). Eye tracking and pupillometry are indicators of dissociable latent decision processes. *Journal of Experimental Psychology: General*, *143*(4), 1476–1488. <https://doi.org/10.1037/a0035813>
- Chabert, T., Colin, A., Aubin, T., Shacks, V., Bourquin, S. L., Elsey, R. M., Acosta, J. G., & Mathevon, N. (2015). Size does matter: Crocodile mothers react more to the voice of smaller offspring. *Scientific Reports*, *5*(1), 15547. <https://doi.org/10.1038/srep15547>
- Dabelsteen, T., McGregor, P. K., Lampe, H. M., Langmore, N. E., & Holland, J. (1998). Quiet song in song birds: An overlooked phenomenon. *Bioacoustics*, *9*(2), 89–105. <https://doi.org/10.1080/09524622.1998.9753385>
- Darwin, C. (1871). The Descent of Man, and Selection in Relation to Sex. In *The Descent of Man, and Selection in Relation to Sex*. Princeton University Press. <https://doi.org/10.1515/9781400820061>
- de Gee, J. W., Knapen, T., & Donner, T. H. (2014). Decision-related pupil dilation reflects upcoming choice and individual bias. *Proceedings of the National Academy of Sciences*, *111*(5), E618–E625. <https://doi.org/10.1073/pnas.1317557111>
- DeGroot, M. H. (1970). *Optimal statistical decisions*. McGraw-Hill.
- Dent, M. L. (1997). Free-field binaural unmasking in budgerigars (*Melopsittacus undulatus*). *Behavioral Neuroscience*, *111*(3), 590–598. <https://doi.org/10.1037/0735-7044.111.3.590>
- Dent, M. L., McClaine, E. M., Best, E., Ozmeral V., Narayan, R., Gallun, F. J., Sen, K., & Shinn-Cunningham, B. G. (2009). Spatial unmasking of birdsong in zebra finches (*Taeniopygia guttata*) and budgerigars (*Melopsittacus undulatus*). *Journal of Comparative Psychology*, *123*(4), 357–367. <https://doi.org/10.1037/a0016898>

- Dooling, R. J., Lohr, B., & Dent, M. L. (2000). Hearing in Birds and Reptiles. In R. J. Dooling, R. R. Fay, & A. N. Popper (Eds.), *Comparative Hearing: Birds and Reptiles* (pp. 308–359). Springer. [https://doi.org/10.1007/978-1-4612-1182-2\\_7](https://doi.org/10.1007/978-1-4612-1182-2_7)
- Douglas, R. H., Collin, S. P., & Corrigan, J. (2002). The eyes of suckermouth armoured catfish (Loricariidae, subfamily Hypostomus): Pupil response, lenticular longitudinal spherical aberration and retinal topography. *Journal of Experimental Biology*, *205*(22), 3425–3433. <https://doi.org/10.1242/jeb.205.22.3425>
- Douglas, R. H., Williamson, R., & Wagner, H.-J. (2005). The pupillary response of cephalopods. *Journal of Experimental Biology*, *208*(2), 261–265. <https://doi.org/10.1242/jeb.01395>
- Duquette, C. A., Loss, S. R., & Hovick, T. J. (2021). A meta-analysis of the influence of anthropogenic noise on terrestrial wildlife communication strategies. *Journal of Applied Ecology*, *58*(6), 1112–1121. <https://doi.org/10.1111/1365-2664.13880>
- Ehret, G. (1992). Categorical perception of mouse-pup ultrasounds in the temporal domain. *Animal Behaviour*, *43*(3), 409–416. [https://doi.org/10.1016/S0003-3472\(05\)80101-0](https://doi.org/10.1016/S0003-3472(05)80101-0)
- Ehret, G. (2006). Common rules of communication sound perception. *Behavior and Neurodynamics for Auditory Communication*, 85–114.
- Ellis, C. J. (1981). The pupillary light reflex in normal subjects. *British Journal of Ophthalmology*, *65*(11), 754–759. <https://doi.org/10.1136/bjo.65.11.754>
- Fabiani, M., Gratton, G., & Federmeier, K. D. (2007). Event-related brain potentials: Methods, theory, and applications. In *Handbook of psychophysiology, 3rd ed* (pp. 85–119). Cambridge University Press. <https://doi.org/10.1017/CBO9780511546396.004>
- Faragó, T., Andics, A., Devecseri, V., Kis, A., Gácsi, M., & Miklósi, Á. (2014). Humans rely on the same rules to assess emotional valence and intensity in conspecific and dog vocalizations. *Biology Letters*, *10*(1), 20130926. <https://doi.org/10.1098/rsbl.2013.0926>
- Faragó, T., Takács, N., Miklósi, A., & Pongracz, P. (2017). Dog growls express various contextual and affective content for human listeners. *Royal Society Open Science*, *4*, 170134. <https://doi.org/10.1098/rsos.170134>
- Filippi, P., Congdon, J. V., Hoang, J., Bowling, D. L., Reber, S. A., Pašukonis, A., Hoeschele, M., Ocklenburg, S., de Boer, B., Sturdy, C. B., Newen, A., & Güntürkün, O. (2017). Humans recognize emotional arousal in vocalizations across all classes of terrestrial vertebrates: Evidence for acoustic universals. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1859), 20170990. <https://doi.org/10.1098/rspb.2017.0990>
- Fischer, J. (1998). Barbary macaques categorize shrill barks into two call types. *Animal Behaviour*, *55*(4), 799–807. <https://doi.org/10.1006/anbe.1997.0663>
- Fischer, J. (2006). Categorical Perception in Animals. *Encyclopedia of Language & Linguistics*. <https://doi.org/10.1016/B0-08-044854-2/00806-3>
- Fitch, W. T., Neubauer, J., & Herzog, H. (2002). Calls out of chaos: The adaptive significance of nonlinear phenomena in mammalian vocal production. *Animal Behaviour*, *63*(3), 407–418. <https://doi.org/10.1006/anbe.2001.1912>
- Footo, A. D., Osborne, R. W., & Hoelzel, A. R. (2004). Whale-call response to masking boat noise. *Nature*, *428*(6986), 910–910. <https://doi.org/10.1038/428910a>



- Freyman, R. L., Balakrishnan, U., & Helfer, K. S. (2001). Spatial release from informational masking in speech recognition. *The Journal of the Acoustical Society of America*, *109*(5), 2112–2122.
- Friedman, D., Hakerem, G., Sutton, S., & Fleiss, J. L. (1973). Effect of stimulus uncertainty on the pupillary dilation response and the vertex evoked potential. *Electroencephalography and Clinical Neurophysiology*, *34*(5), 475–484. [https://doi.org/10.1016/0013-4694\(73\)90065-5](https://doi.org/10.1016/0013-4694(73)90065-5)
- Furuyama, T., Kobayasi, K. I., & Riquimaroux, H. (2017). Acoustic characteristics used by Japanese macaques for individual discrimination. *Journal of Experimental Biology*, *220*(19), 3571–3578. <https://doi.org/10.1242/jeb.154765>
- Garrick, L. D., & Lang, J. W. (1977). Social Signals and Behaviors of Adult Alligators and Crocodiles. *American Zoologist*, *17*(1), 225–239. <https://doi.org/10.1093/icb/17.1.225>
- Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models. *Statistics and Computing*, *24*(6), 997–1016. <https://doi.org/10.1007/s11222-013-9416-2>
- Gelman, A., Rubin, D. B., & others. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457–472. <https://doi.org/10.1214/ss/1177011136>
- Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, *10*(2), 252–269. <https://doi.org/10.3758/CABN.10.2.252>
- Gleich, O., & Manley, G. A. (2000). The Hearing Organ of Birds and Crocodylia. In R. J. Dooling, R. R. Fay, & A. N. Popper (Eds.), *Comparative Hearing: Birds and Reptiles* (pp. 70–138). Springer. [https://doi.org/10.1007/978-1-4612-1182-2\\_3](https://doi.org/10.1007/978-1-4612-1182-2_3)
- Gomes, D. G. E., Francis, C. D., & Barber, J. R. (2021). Using the Past to Understand the Present: Coping with Natural and Anthropogenic Noise. *BioScience*, *71*(3), 223–234. <https://doi.org/10.1093/biosci/biaa161>
- Green, P. A., Brandley, N. C., & Nowicki, S. (2020). Categorical perception in animal communication and decision-making. *Behavioral Ecology*, *31*(4), 859–867. <https://doi.org/10.1093/beheco/araa004>
- Greene, N. T., Anbuhl, K. L., Ferber, A. T., DeGuzman, M., Allen, P. D., & Tollin, D. J. (2018). Spatial hearing ability of the pigmented Guinea pig (*Cavia porcellus*): Minimum audible angle and spatial release from masking in azimuth. *Hearing Research*, *365*, 62–76. <https://doi.org/10.1016/j.heares.2018.04.011>
- Grigg, G. C., & Kirshner, D. (2015). *Biology and evolution of crocodylians*. Comstock Publishing Associates a division of Cornell University Press.
- Guilford, T., & Dawkins, M. S. (1991). Receiver psychology and the evolution of animal signals. *Animal Behaviour*, *42*(1), 1–14. [https://doi.org/10.1016/S0003-3472\(05\)80600-1](https://doi.org/10.1016/S0003-3472(05)80600-1)
- Harnad, S. (1987). *Categorical perception: The groundwork of cognition* (pp. x, 599). Cambridge University Press.

- Herzog, H. A., & Burghardt, G. M. (1977). Vocalization in Juvenile Crocodilians. *Zeitschrift Für Tierpsychologie*, 44(3), 294–304. <https://doi.org/10.1111/j.1439-0310.1977.tb00997.x>
- Higgs, D., Brittan-Powell, E., Soares, D., Souza, M., Carr, C., Dooling, R., & Popper, A. (2002). Amphibious auditory responses of the American alligator (*Alligator mississippiensis*). *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology*, 188, 217–223. <https://doi.org/10.1007/s00359-002-0296-8>
- Hine, J. E., Martin, R. L., & Moore, D. R. (1994). Free-field binaural unmasking in ferrets. *Behavioral Neuroscience*, 108(1), 196–205. <https://doi.org/10.1037/0735-7044.108.1.196>
- Holt, M. M., & Schusterman, R. J. (2007). Spatial release from masking of aerial tones in pinnipeds. *The Journal of the Acoustical Society of America*, 121(2), 1219–1225. <https://doi.org/10.1121/1.2404929>
- Hugall, A. F., Foster, R., & Lee, M. S. Y. (2007). Calibration choice, rate smoothing, and the pattern of tetrapod diversification according to the long nuclear gene RAG-1. *Systematic Biology*, 56(4), 543–563. <https://doi.org/10.1080/10635150701477825>
- Hulse, S. H. (2002). Auditory scene analysis in animal communication. In *Advances in the Study of Behavior* (Vol. 31, pp. 163–200). Academic Press. [https://doi.org/10.1016/S0065-3454\(02\)80008-0](https://doi.org/10.1016/S0065-3454(02)80008-0)
- Hunt, R. H., & Watanabe, M. E. (1982). Observations on Maternal Behavior of the American Alligator, *Alligator mississippiensis*. *Journal of Herpetology*, 16(3), 235–239. <https://doi.org/10.2307/1563716>
- Hutton, J. (1989). Movements, Home Range, Dispersal and the Separation of Size Classes in Nile Crocodiles. *American Zoologist*, 29(3), 1033–1049.
- Jones, G. L., & Litovsky, R. Y. (2011). A cocktail party model of spatial release from masking by both noise and speech interferers. *The Journal of the Acoustical Society of America*, 130(3), 1463–1474. <https://doi.org/10.1121/1.3613928>
- Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus, Colliculi, and Cingulate Cortex. *Neuron*, 89(1), 221–234. <https://doi.org/10.1016/j.neuron.2015.11.028>
- Kang, O. E., Huffer, K. E., & Wheatley, T. P. (2014). Pupil Dilation Dynamics Track Attention to High-Level Information. *PLOS ONE*, 9(8), e102463. <https://doi.org/10.1371/journal.pone.0102463>
- Kawahara, H. (2006). STRAIGHT, exploitation of the other aspect of VOCODER: Perceptually isomorphic decomposition of speech sounds. *Acoustical Science and Technology*, 27(6), 349–353. <https://doi.org/10.1250/ast.27.349>
- Kawahara, H., Masuda-Katsuse, I., & de Cheveigné, A. (1999). Restructuring speech representations using a pitch-adaptive time–frequency smoothing and an instantaneous-frequency-based F0 extraction: Possible role of a repetitive structure in sounds | Speech files available. See <http://www.elsevier.nl/locate/specom1>. *Speech Communication*, 27(3), 187–207. [https://doi.org/10.1016/S0167-6393\(98\)00085-5](https://doi.org/10.1016/S0167-6393(98)00085-5)
- Kelly, T., Reby, D., Levréro, F., Keenan, S., Gustafsson, E., Koutseff, A., & Mathevon, N. (2017). Adult human perception of distress in the cries of bonobo, chimpanzee, and



- human infants. *Biological Journal of the Linnean Society*, 120(4), 919–930.  
<https://doi.org/10.1093/biolinnean/blw016>
- Kettler, L., & Carr, C. E. (2019). Neural Maps of Interaural Time Difference in the American Alligator: A Stable Feature in Modern Archosaurs. *Journal of Neuroscience*, 39(20), 3882–3896. <https://doi.org/10.1523/JNEUROSCI.2989-18.2019>
- Klinke, R., & Pause, M. (1980). Discharge properties of primary auditory fibres in caiman crocodilus: Comparisons and contrasts to the mammalian auditory nerve. *Experimental Brain Research*, 38(2), 137–150. <https://doi.org/10.1007/BF00236735>
- Kloosterman, N. A., Meindertsma, T., van Loon, A. M., Lamme, V. A. F., Bonnef, Y. S., & Donner, T. H. (2015). Pupil size tracks perceptual content and surprise. *European Journal of Neuroscience*, 41(8), 1068–1078. <https://doi.org/10.1111/ejn.12859>
- Kluender, K. R., Diehl, R. L., & Killeen, P. R. (1987). Japanese quail can learn phonetic categories. *Science (New York, N.Y.)*, 237(4819), 1195–1197.  
<https://doi.org/10.1126/science.3629235>
- Knudsen, E. I., & Konishi, M. (1979). Mechanisms of sound localization in the barn owl (*Tyto alba*). *Journal of Comparative Physiology ? A*, 133(1), 13–21.  
<https://doi.org/10.1007/BF00663106>
- Koutseff, A., Reby, D., Martin, O., Levrero, F., Patural, H., & Mathevon, N. (2018). The acoustic space of pain: Cries as indicators of distress recovering dynamics in pre-verbal infants. *Bioacoustics*, 27(4), 313–325.  
<https://doi.org/10.1080/09524622.2017.1344931>
- Kruschke, J. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press.
- Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review*, 25(1), 178–206. <https://doi.org/10.3758/s13423-016-1221-4>
- Kuhl, P. K., & Miller, J. D. (1978). Speech perception by the chinchilla: Identification functions for synthetic VOT stimuli. *The Journal of the Acoustical Society of America*, 63(3), 905–917. <https://doi.org/10.1121/1.381770>
- Lang, J., Whitaker, R., & Andrews, H. (1986). Male parental care in mugger crocodiles. *National Geographic Research*, 2, 519–525.
- Laukka, P., Juslin, P., & Bresin, R. (2005). A dimensional approach to vocal expression of emotion. *Cognition and Emotion*, 19(5), 633–653.  
<https://doi.org/10.1080/02699930441000445>
- Laurijs, K. A., Briefer, E. F., Reimert, I., & Webb, L. E. (2021). Vocalisations in farm animals: A step towards positive welfare assessment. *Applied Animal Behaviour Science*, 236, 105264. <https://doi.org/10.1016/j.applanim.2021.105264>
- Lavin, C., San Martín, R., & Rosales Jubal, E. (2014). Pupil dilation signals uncertainty and surprise in a learning gambling task. *Frontiers in Behavioral Neuroscience*, 7.  
<https://www.frontiersin.org/articles/10.3389/fnbeh.2013.00218>
- Lee, D. S. (1968). Possible communication between eggs of the American alligator. *Herpetologica*.

- Lee, N., Christensen-Dalsgaard, J., White, L. A., Schrode, K. M., & Bee, M. A. (2021). Lung mediated auditory contrast enhancement improves the Signal-to-noise ratio for communication in frogs. *Current Biology*, *31*(7), 1488-1498.e4. <https://doi.org/10.1016/j.cub.2021.01.048>
- Lee, N., & Mason, A. C. (2017). How spatial release from masking may fail to function in a highly directional auditory system. *ELife*, *6*. <https://doi.org/10.7554/eLife.20731>
- Lengagne, T., Aubin, T., Lauga, J., & Jouventin, P. (1999). How do king penguins (*Aptenodytes patagonicus*) apply the mathematical theory of information to communicate in windy conditions? *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *266*(1429), 1623–1628. <https://doi.org/10.1098/rspb.1999.0824>
- Li, T., & Howland, H. C. (1999). A true neuronal consensual pupillary reflex in chicks. *Vision Research*, *39*(5), 897–900. [https://doi.org/10.1016/S0042-6989\(98\)00197-7](https://doi.org/10.1016/S0042-6989(98)00197-7)
- Lingle, S., & Riede, T. (2014). Deer Mothers Are Sensitive to Infant Distress Vocalizations of Diverse Mammalian Species. *The American Naturalist*, *184*(4), 510–522. <https://doi.org/10.1086/677677>
- Lingle, S., WYMAN, M., Kotrba, R., TEICHROEB, L., & Romanow, C. (2012). What makes a cry a cry? A review of infant distress vocalizations. *Current Zoology*, *58*, 698–726. <https://doi.org/10.1093/czoolo/58.5.698>
- Litovsky, R. (2012). Spatial Release from Masking. *Acoustics Today*, *8*, 18. <https://doi.org/10.1121/1.4729575>
- Lohr, B., Wright, T. F., & Dooling, R. J. (2003). Detection and discrimination of natural calls in masking noise by birds: Estimating the active space of a signal. *Animal Behaviour*, *65*(4), 763–777. <https://doi.org/10.1006/anbe.2003.2093>
- Manabe, K., Sadr, E. I., & Dooling, R. J. (1998). Control of vocal intensity in budgerigars (*Melopsittacus undulatus*): Differential reinforcement of vocal intensity and the Lombard effect. *The Journal of the Acoustical Society of America*, *103*(2), 1190–1198. <https://doi.org/10.1121/1.421227>
- Manley, G. A. (1970). Frequency sensitivity of auditory neurons in the Caiman cochlear nucleus. *Zeitschrift Für Vergleichende Physiologie*, *66*(3), 251–256. <https://doi.org/10.1007/BF00297828>
- Manteuffel, G., Puppe, B., & Schön, P. C. (2004). Vocalization of farm animals as a measure of welfare. *Applied Animal Behaviour Science*, *88*(1), 163–182. <https://doi.org/10.1016/j.applanim.2004.02.012>
- Maruščáková, I. L., Linhart, P., Ratcliffe, V. F., Tallet, C., Reby, D., & Špinka, M. (2015). Humans (*Homo sapiens*) judge the emotional content of piglet (*Sus scrofa domestica*) calls based on simple acoustic parameters, not personality, empathy, nor attitude toward animals. *Journal of Comparative Psychology*, *129*(2), 121–131. <https://doi.org/10.1037/a0038870>
- Marx, A., Lenkei, R., Pérez Fraga, P., Bakos, V., Kubinyi, E., & Faragó, T. (2021). Occurrences of non-linear phenomena and vocal harshness in dog whines as indicators of stress and ageing. *Scientific Reports*, *11*(1), 4468. <https://doi.org/10.1038/s41598-021-83614-1>



- Massenet, M., Anikin, A., Pisanski, K., Reynaud, K., Mathevon, N., & Reby, D. (2022). Nonlinear vocal phenomena affect human perceptions of distress, size and dominance in puppy whines. *Proceedings of the Royal Society B: Biological Sciences*, 289(1973), 20220429. <https://doi.org/10.1098/rspb.2022.0429>
- Mathevon, N., & Aubin, T. (1997). Reaction to conspecific degraded song by the wren Troglodytes troglodytes: Territorial response and choice of song post. *Behavioural Processes*, 39(1), 77–84. [https://doi.org/10.1016/S0376-6357\(96\)00046-0](https://doi.org/10.1016/S0376-6357(96)00046-0)
- Mathevon, N., & Aubin, T. (2020). *Coding Strategies in Vertebrate Acoustic Communication*. Springer. <https://doi.org/10.1007/978-3-030-39200-0>
- Mathevon, N., Aubin, T., Vielliard, J., da Silva, M.-L., Sèbe, F., & Boscolo, D. (2008). Singing in the rain Forest: How a tropical bird song transfers information. *PLoS ONE*, 3(2), e1580. <https://doi.org/10.1371/journal.pone.0001580>
- Mathevon, N., Dabelsteen, T., & Blumenrath, S. H. (2005). Are high perches in the blackcap *Sylvia atricapilla* song or listening posts? A sound transmission study. *The Journal of the Acoustical Society of America*, 117(1), 442–449. <https://doi.org/10.1121/1.1828805>
- Mathôt, S. (2018). Pupillometry: Psychology, Physiology, and Function. *Journal of Cognition*, 1(1), 16. <https://doi.org/10.5334/joc.18>
- May, B., Moody, D. B., & Stebbins, W. C. (1989). Categorical perception of conspecific communication sounds by Japanese macaques, *Macaca fuscata*. *The Journal of the Acoustical Society of America*, 85(2), 837–847. <https://doi.org/10.1121/1.397555>
- McComb, K., Taylor, A. M., Wilson, C., & Charlton, B. D. (2009). The cry embedded within the purr. *Current Biology*, 19(13), R507–R508. <https://doi.org/10.1016/j.cub.2009.05.033>
- Mendl, M., Burman, O. H. P., & Paul, E. S. (2010). An integrative and functional framework for the study of animal emotion and mood. *Proceedings of the Royal Society B: Biological Sciences*, 277(1696), 2895–2904. <https://doi.org/10.1098/rspb.2010.0303>
- Miller, C. T., & Bee, M. A. (2012). Receiver psychology turns 20: Is it time for a broader approach? *Animal Behaviour*, 83(2), 331–343. <https://doi.org/10.1016/j.anbehav.2011.11.025>
- Montes-Lourido, P., Kar, M., Kumbam, I., & Sadagopan, S. (2021). Pupillometry as a reliable metric of auditory detection and discrimination across diverse stimulus paradigms in animal models. *Scientific Reports*, 11(1), 3108. <https://doi.org/10.1038/s41598-021-82340-y>
- Morse, P. A., & Snowdon, C. T. (1975). An investigation of categorical speech discrimination by rhesus monkeys. *Perception & Psychophysics*, 17(1), 9–16. <https://doi.org/10.3758/BF03203991>
- Morton, E. S. (1977). On the Occurrence and Significance of Motivation-Structural Rules in Some Bird and Mammal Sounds. *The American Naturalist*, 111(981), 855–869. <https://doi.org/10.1086/283219>
- Murphy, P. R., Vandekerckhove, J., & Nieuwenhuis, S. (2014). Pupil-Linked Arousal Determines Variability in Perceptual Decision Making. *PLOS Computational Biology*, 10(9), e1003854. <https://doi.org/10.1371/journal.pcbi.1003854>

- Murray, C. M., Russo, P., Zorrilla, A., & McMahan, C. D. (2019). Divergent Morphology among Populations of the New Guinea Crocodile, *Crocodylus novaeguineae* (Schmidt, 1928): Diagnosis of an Independent Lineage and Description of a New Species. *Copeia*, *107*(3), 517–523. <https://doi.org/10.1643/CG-19-240>
- Nagloo, N., Collin, S. P., Hemmi, J. M., & Hart, N. S. (2016). Spatial resolving power and spectral sensitivity of the saltwater crocodile, *Crocodylus porosus*, and the freshwater crocodile, *Crocodylus johnstoni*. *Journal of Experimental Biology*, *219*(9), 1394–1404. <https://doi.org/10.1242/jeb.135673>
- Naifeh, K. H., Huggins, S. E., Hoff, H. E., Hugg, T. W., & Norton, R. E. (1970). Respiratory patterns in crocodylian reptiles. *Respiration Physiology*, *9*(1), 21–42. [https://doi.org/10.1016/0034-5687\(70\)90003-4](https://doi.org/10.1016/0034-5687(70)90003-4)
- Nelson, D. A., & Marler, P. (1989). Categorical perception of a natural stimulus continuum: Birdsong. *Science (New York, N.Y.)*, *244*(4907), 976–978. <https://doi.org/10.1126/science.2727689>
- Nemeth, E., Dabelsteen, T., Pedersen, S. B., & Winkler, H. (2006). Rainforests as concert halls for birds: Are reverberations improving sound transmission of long song elements? *The Journal of the Acoustical Society of America*, *119*(1), 620–626. <https://doi.org/10.1121/1.2139072>
- Nesbitt, S. J. (2011). *The early evolution of archosaurs: Relationships and the origin of major clades*. (*Bulletin of the American Museum of Natural History*, no. 352). <https://digitallibrary.amnh.org/handle/2246/6112>
- Nieuwenhuis, S., De Geus, E. J., & Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology*, *48*(2), 162–175. <https://doi.org/10.1111/j.1469-8986.2010.01057.x>
- Nityananda, V., & Bee, M. A. (2012). Spatial release from masking in a free-field source identification task by gray treefrogs. *Hearing Research*, *285*(1–2), 86–97. <https://doi.org/10.1016/j.heares.2012.01.003>
- Papet, L., Grimault, N., Boyer, N., & Mathevon, N. (2019). Influence of head morphology and natural postures on sound localization cues in crocodylians. *Royal Society Open Science*, *6*(7), 190423. <https://doi.org/10.1098/rsos.190423>
- Papet, L., Raymond, M., Boyer, N., Mathevon, N., & Grimault, N. (2020). Crocodiles use both interaural level differences and interaural time differences to locate a sound source. *The Journal of the Acoustical Society of America*, *148*(4), EL307–EL313. <https://doi.org/10.1121/10.0001979>
- Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex* (pp. xv, 430). Oxford Univ. Press.
- Pongrácz, P., Molnár, C., & Miklósi, Á. (2006). Acoustic parameters of dog barks carry emotional information for humans. *Applied Animal Behaviour Science*, *100*(3), 228–240. <https://doi.org/10.1016/j.aplanim.2005.12.004>
- Popov, V. V., Supin, A. Ya., Gvozdeva, A. P., Nechaev, D. I., Tarakanov, M. B., & Sysueva, E. V. (2020). Spatial release from masking in a bottlenose dolphin *Tursiops truncatus*. *The Journal of the Acoustical Society of America*, *147*(3), 1719–1726. <https://doi.org/10.1121/10.0000909>





- Prato-Previde, E., Cannas, S., Palestini, C., Ingrassia, S., Battini, M., Ludovico, L. A., Ntalampiras, S., Presti, G., & Mattiello, S. (2020). What's in a Meow? A Study on Human Classification and Interpretation of Domestic Cat Vocalizations. *Animals: An Open Access Journal from MDPI*, *10*(12), E2390. <https://doi.org/10.3390/ani10122390>
- Preuschoff, K., 't Hart, B., & Einhauser, W. (2011). Pupil Dilation Signals Surprise: Evidence for Noradrenaline's Role in Decision Making. *Frontiers in Neuroscience*, *5*. <https://www.frontiersin.org/articles/10.3389/fnins.2011.00115>
- Puig-Diví, A., Escalona-Marfil, C., Padullés-Riu, J. M., Busquets, A., Padullés-Chando, X., & Marcos-Ruiz, D. (2019). Validity and reliability of the Kinovea program in obtaining angles and distances using coordinates in 4 perspectives. *PLOS ONE*, *14*(6), e0216448. <https://doi.org/10.1371/journal.pone.0216448>
- Qiyuan, J., Richer, F., Wagoner, B. L., & Beatty, J. (1985). The Pupil and Stimulus Probability. *Psychophysiology*, *22*(5), 530–534. <https://doi.org/10.1111/j.1469-8986.1985.tb01645.x>
- Raine, J., Pisanski, K., Simner, J., & Reby, D. (2019). Vocal communication of simulated pain. *Bioacoustics*, *28*(5), 404–426. <https://doi.org/10.1080/09524622.2018.1463295>
- Ratnam, R., & Feng, A. S. (1998). Detection of Auditory Signals by Frog Inferior Collicular Neurons in the Presence of Spatially Separated Noise. *Journal of Neurophysiology*, *80*(6), 2848–2859. <https://doi.org/10.1152/jn.1998.80.6.2848>
- Reber, S. A. (2020). Crocodilians Are Promising Intermediate Model Organisms for Comparative Perception Research. *Comparative Cognition & Behavior Reviews*, *15*, 111–129. <https://doi.org/10.3819/CCBR.2020.150004>
- Reese, A. M. (1945). The laryngeal region of alligator mississippiensis. *The Anatomical Record*, *92*(3), 273–277. <https://doi.org/10.1002/ar.1090920308>
- Reinhard, G., & Lachnit, H. (2002). The effect of stimulus probability on pupillary response as an indicator of cognitive processing in human learning and categorization. *Biological Psychology*, *60*(2–3), 199–215. [https://doi.org/10.1016/S0301-0511\(02\)00031-5](https://doi.org/10.1016/S0301-0511(02)00031-5)
- Rendall, D., Notman, H., & Owren, M. J. (2009). Asymmetries in the individual distinctiveness and maternal recognition of infant contact calls and distress screams in baboons. *The Journal of the Acoustical Society of America*, *125*(3), 1792–1805. <https://doi.org/10.1121/1.3068453>
- Richer, F., & Beatty, J. (1987). Contrasting Effects of Response Uncertainty on the Task-Evoked Pupillary Response and Reaction Time. *Psychophysiology*, *24*(3), 258–262. <https://doi.org/10.1111/j.1469-8986.1987.tb00291.x>
- Rocchi, F., Dylla, M. E., Bohlen, P. A., & Ramachandran, R. (2017). Spatial and temporal disparity in signals and maskers affects signal detection in non-human primates. *Hearing Research*, *344*, 1–12. <https://doi.org/10.1016/j.heares.2016.10.013>
- Root-Gutteridge, H., Ratcliffe, V. F., Neumann, J., Timarchi, L., Yeung, C., Korzeniowska, A. T., Mathevon, N., & Reby, D. (2021). Effect of pitch range on dogs' response to conspecific vs. Heterospecific distress cries. *Scientific Reports*, *11*(1), 19723. <https://doi.org/10.1038/s41598-021-98967-w>

- Rowe, C. (2013). Receiver psychology: A receiver's perspective. *Animal Behaviour*, 85(3), 517–523. <https://doi.org/10.1016/j.anbehav.2013.01.004>
- Russell, A. P., & Bauer, A. M. (2021). Vocalization by extant nonavian reptiles: A synthetic overview of phonation and the vocal apparatus. *The Anatomical Record*, 304(7), 1478–1528. <https://doi.org/10.1002/ar.24553>
- Saberi, K., Dostal, L., Sadralodabai, T., Bull, V., & Perrott, D. R. (1991). Free-field release from masking. *The Journal of the Acoustical Society of America*, 90(3), 1355–1370. <https://doi.org/10.1121/1.401927>
- Saunders, J. C., Duncan, R. K., Doan, D. E., & Werner, Y. L. (2000). The Middle Ear of Reptiles and Birds. In R. J. Dooling, R. R. Fay, & A. N. Popper (Eds.), *Comparative Hearing: Birds and Reptiles* (pp. 13–69). Springer. [https://doi.org/10.1007/978-1-4612-1182-2\\_2](https://doi.org/10.1007/978-1-4612-1182-2_2)
- Schaeffel, F., & Wagner, H. (1992). Barn owls have symmetrical accommodation in both eyes, but independent pupillary responses to light. *Vision Research*, 32(6), 1149–1155. [https://doi.org/10.1016/0042-6989\(92\)90016-C](https://doi.org/10.1016/0042-6989(92)90016-C)
- Scheumann, M., Hasting, A. S., Kotz, S. A., & Zimmermann, E. (2014). The Voice of Emotion across Species: How Do Human Listeners Recognize Animals' Affective States? *PLOS ONE*, 9(3), e91192. <https://doi.org/10.1371/journal.pone.0091192>
- Schmidt, A. K. D., & Romer, H. (2011). Solutions to the Cocktail Party Problem in Insects: Selective Filters, Spatial Release from Masking and Gain Control in Tropical Crickets. *PLoS ONE*, 6(12), e28593. <https://doi.org/10.1371/journal.pone.0028593>
- Schnupp, J. W., & Carr, C. E. (2009). On hearing with more than one ear: Lessons from evolution. *Nature Neuroscience*, 12(6), 692. <https://doi.org/10.1038/nn.2325>
- Shimizu, T., Patton, T., Szafranski, G., & Butler, A. (2009). *Evolution of the Visual System in Reptiles and Birds* (pp. 1466–1472). [https://doi.org/10.1007/978-3-540-29678-2\\_3179](https://doi.org/10.1007/978-3-540-29678-2_3179)
- Sicuro, F. L., Iack-Ximenes, G. E., Wogel, H., & Bilate, M. (2013). Vocal patterns of adult females and juveniles Caiman yacare (Crocodylia: *Alligatoridae*) in Brazilian Pantanal wetland. *Rev. Biol. Trop.*, 61, 14.
- Sillman, A. J., Ronan, S. J., & Loew, E. R. (1991). Histology and microspectrophotometry of the photoreceptors of a crocodylian, *Alligator mississippiensis*. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 243(1306), 93–98. <https://doi.org/10.1098/rspb.1991.0016>
- Sinnott, J. M., Beecher, M. D., Moody, D. B., & Stebbins, W. C. (1976). Speech sound discrimination by monkeys and humans. *The Journal of the Acoustical Society of America*, 60(3), 687–695. <https://doi.org/10.1121/1.381140>
- Slabbekoorn, H., & Peet, M. (2003). Birds sing at a higher pitch in urban noise. *Nature*, 424(6946), 267–267. <https://doi.org/10.1038/424267a>
- Somaweera, R., Brien, M., & Shine, R. (2013). The Role of Predation in Shaping Crocodylian Natural History. *Herpetological Monographs*, 27(1), 23. <https://doi.org/10.1655/HERPMONOGRAPHS-D-11-00001>
- Somaweera, R., & Shine, R. (2012). Australian Freshwater Crocodiles (*Crocodylus johnstoni*) Transport Their Hatchlings to the Water. *Journal of Herpetology*, 46(3), 407–411. <https://doi.org/10.1670/11-056>
- Stevenson, C. (2019). *Crocodiles of the World*. New Holland Publishers.



- Stoeger, A. S., Charlton, B. D., Kratochvil, H., & Fitch, W. T. (2011). Vocal cues indicate level of arousal in infant African elephant roars. *The Journal of the Acoustical Society of America*, *130*(3), 1700–1710. <https://doi.org/10.1121/1.3605538>
- Sümer, S., Denzinger, A., & Schnitzler, H.-U. (2009). Spatial unmasking in the echolocating Big Brown Bat, *Eptesicus fuscus*. *Journal of Comparative Physiology A*, *195*(5), 463–472. <https://doi.org/10.1007/s00359-009-0424-9>
- Tallet, C., Špinka, M., Maruščáková, I., & Šimeček, P. (2010). Human perception of vocalizations of domestic piglets and modulation by experience with domestic pigs (*Sus scrofa*). *Journal of Comparative Psychology*, *124*(1), 81–91. <https://doi.org/10.1037/a0017354>
- Teichroeb, L. J., Riede, T., Kotrba, R., & Lingle, S. (2013). Fundamental frequency is key to response of female deer to juvenile distress calls. *Behavioural Processes*, *92*, 15–23. <https://doi.org/10.1016/j.beproc.2012.09.011>
- Thévenet, J., Papet, L., Campos, Z., Greenfield, M., Boyer, N., Grimault, N., & Mathevon, N. (2022a). Data from: Spatial release from masking in crocodilians. <https://doi.org/10.5281/zenodo.5971364>
- Thévenet, J., Papet, L., Campos, Z., Greenfield, M., Boyer, N., Grimault, N., & Mathevon, N. (2022b). Spatial release from masking in crocodilians. *Communications Biology*, *5*(1), 1–14. <https://doi.org/10.1038/s42003-022-03799-7>
- Todd, N. P. M. (2007). Estimated source intensity and active space of the American alligator (*Alligator mississippiensis*) vocal display. *The Journal of the Acoustical Society of America*, *122*(5), 2906–2915. <https://doi.org/10.1121/1.2785811>
- Urai, A. E., Braun, A., & Donner, T. H. (2017). Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nature Communications*, *8*(1), 14637. <https://doi.org/10.1038/ncomms14637>
- Vergne, A., Avril, A., Martin, S., & Mathevon, N. (2007). Parent-offspring communication in the Nile crocodile *Crocodylus niloticus*: Do newborns' calls show an individual signature? *Die Naturwissenschaften*, *94*, 49–54. <https://doi.org/10.1007/s00114-006-0156-4>
- Vergne, A. L., Aubin, T., Martin, S., & Mathevon, N. (2012). Acoustic communication in crocodilians: Information encoding and species specificity of juvenile calls. *Animal Cognition*, *15*(6), 1095–1109. <https://doi.org/10.1007/s10071-012-0533-7>
- Vergne, A. L., Aubin, T., Taylor, P., & Mathevon, N. (2011). Acoustic signals of baby black caimans. *Zoology*, *114*(6), 313–320. <https://doi.org/10.1016/j.zool.2011.07.003>
- Vergne, A. L., & Mathevon, N. (2008). Crocodile egg sounds signal hatching time. *Current Biology*, *18*(12), R513–R514. <https://doi.org/10.1016/j.cub.2008.04.011>
- Vergne, A. L., Pritz, M. B., & Mathevon, N. (2009). Acoustic communication in crocodilians: From behaviour to brain. *Biological Reviews*, *84*(3), 391–411. <https://doi.org/10.1111/j.1469-185X.2009.00079.x>
- Vergne, A., Pritz, M., & Mathevon, N. (2009). Acoustic communication in crocodilians: From behaviour to brain. *Biological Reviews of the Cambridge Philosophical Society*, *84*, 391–411. <https://doi.org/10.1111/j.1469-185X.2009.00079.x>
- Villain, A. S., Fernandez, M. S. A., Bouchut, C., Soula, H. A., & Vignal, C. (2016). Songbird mates change their call structure and intrapair communication at the nest in response

- to environmental noise. *Animal Behaviour*, *116*, 113–129.  
<https://doi.org/10.1016/j.anbehav.2016.03.009>
- Vliet, K. A. (1989). Social Displays of the American Alligator (*Alligator mississippiensis*)1. *American Zoologist*, *29*(3), 1019–1031. <https://doi.org/10.1093/icb/29.3.1019>
- Waal, F. B. M. D. (1988). The Communicative Repertoire of Captive Bonobos (*Pan Paniscus*), Compared To That of Chimpanzees. *Behaviour*, *106*(3–4), 183–251.  
<https://doi.org/10.1163/156853988X00269>
- Wakeford, O. S., & Robinson, D. E. (1974). Detection of binaurally masked tones by the cat. *The Journal of the Acoustical Society of America*, *56*(3), 952–956.  
<https://doi.org/10.1121/1.1903354>
- Wang, X., Wang, D., Wu, X., Wang, R., & Wang, C. (2007). Acoustic signals of Chinese alligators (*Alligator sinensis*): Social communication. *The Journal of the Acoustical Society of America*, *121*(5), 2984–2989. <https://doi.org/10.1121/1.2714910>
- Weary, D. M. (1990). Categorization of song notes in great tits: Which acoustic features are used and why? *Animal Behaviour*, *39*(3), 450–457. [https://doi.org/10.1016/S0003-3472\(05\)80408-7](https://doi.org/10.1016/S0003-3472(05)80408-7)
- Weber, E. H. (1834). *De Pulsu, resorptione, auditu et tactu: Annotationes anatomicae et physiologicae*... C.F. Koehler.
- Wetzel, W., Wagner, T., Ohl, F. W., & Scheich, H. (1998). Categorical discrimination of direction in frequency-modulated tones by Mongolian gerbils. *Behavioural Brain Research*, *91*(1), 29–39. [https://doi.org/10.1016/S0166-4328\(97\)00099-5](https://doi.org/10.1016/S0166-4328(97)00099-5)
- Wever, E. G., & Vernon, J. A. (1957). Auditory responses in the spectacled caiman. *Journal of Cellular and Comparative Physiology*, *50*(2), 333–339.  
<https://doi.org/10.1002/jcp.1030500213>
- Wierda, S. M., van Rijn, H., Taatgen, N. A., & Martens, S. (2012). Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. *Proceedings of the National Academy of Sciences*, *109*(22), 8456–8460.  
<https://doi.org/10.1073/pnas.1201858109>
- Wilden, I., Herzel, H., Peters, G., & Tembrock, G. (1998). Subharmonics, Biphonation, and Deterministic Chaos in Mammal Vocalization. *Bioacoustics*, *9*(3), 171–196.  
<https://doi.org/10.1080/09524622.1998.9753394>
- Witmer, L. M., Ridgely, R. C., Dufeu, D. L., & Semones, M. C. (2008). Using CT to Peer into the Past: 3D Visualization of the Brain and Ear Regions of Birds, Crocodiles, and Nonavian Dinosaurs. In H. Endo & R. Frey (Eds.), *Anatomical Imaging: Towards a New Morphology* (pp. 67–87). Springer Japan. [https://doi.org/10.1007/978-4-431-76933-0\\_6](https://doi.org/10.1007/978-4-431-76933-0_6)
- Wood, C. C. (1976). Discriminability, response bias, and phoneme categories in discrimination of voice onset time. *The Journal of the Acoustical Society of America*, *60*(6), 1381–1389. <https://doi.org/10.1121/1.381231>
- Wytenbach, R. A., May, M. L., & Hoy, R. R. (1996). Categorical Perception of Sound Frequency by Crickets. *Science*, *273*(5281), 1542–1544.
- Zentall, T. R., Galizio, M., & Critchfield, T. S. (2002). Categorization, Concept Learning, and Behavior Analysis: An Introduction. *Journal of the Experimental Analysis of Behavior*, *78*(3), 237–248. <https://doi.org/10.1901/jeab.2002.78-237>





# Annex

**Peripheral publication: Thévenet, J., Grimault, N., Fonseca, P., & Mathevon, N. (2022). Voice-mediated interactions in a megaherbivore.**

**Current Biology, 32(2), R70–R71.**

**<https://doi.org/10.1016/j.cub.2021.12.017>**





## Correspondence

# Voice-mediated interactions in a megaherbivore

Julie Thévenet<sup>1,2</sup>, Nicolas Grimault<sup>2</sup>,  
Paulo Fonseca<sup>3,5,\*</sup>,  
and Nicolas Mathevon<sup>1,4,5,6,\*</sup>

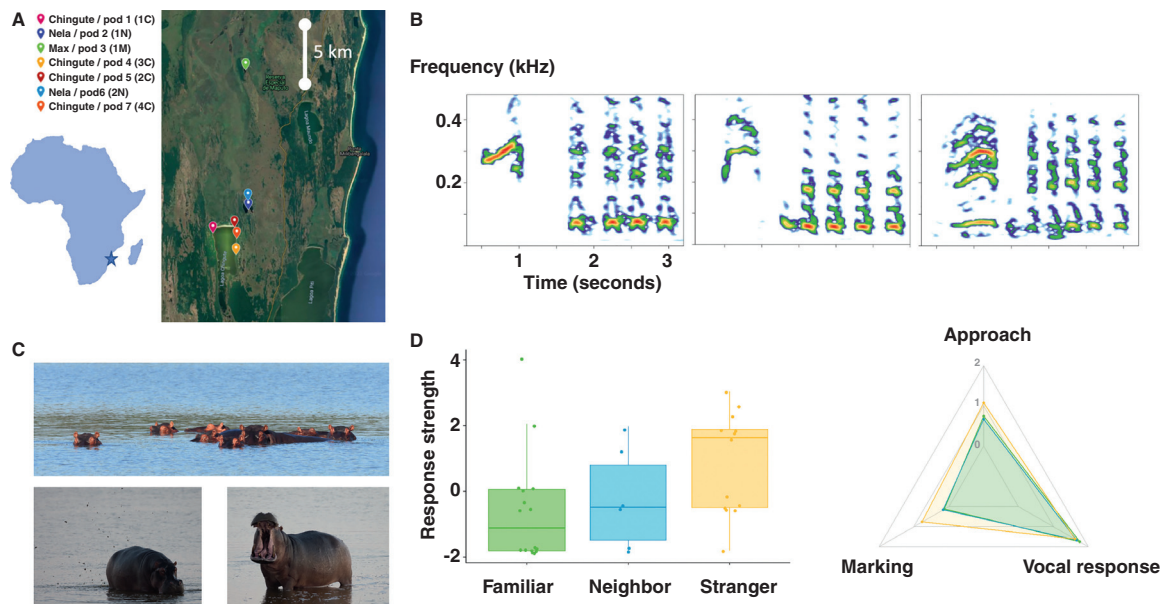
Planet Earth is becoming increasingly difficult for large animal species to inhabit. Yet, these species are of major importance for the functioning of the biosphere and their progressive disappearance is accompanied by profound negative alterations of ecosystems<sup>1</sup> (Supplemental information). To implement effective conservation measures, it is essential to have a detailed knowledge of the biology of these species. Here, we show that the hippopotamus *Hippopotamus amphibius*, an iconic African megaherbivore for which little

is known about social communication, uses vocal recognition to manage relationships between territorial groups. We conducted playback experiments on groups of hippos and observed their response to vocalizations from an individual of the same group (familiar), a group from the same lake (neighbor) and a distant group (stranger). We found that stranger vocalizations induced a stronger behavioral reaction than the other two stimuli. In addition to showing that hippos are able to identify categories of conspecifics based on vocal signatures, our study demonstrates that hippo groups are territorial entities that behave less aggressively toward their neighbors than toward strangers. These new behavioral data suggest that habituation playbacks prior to conservation translocation operations may help reduce the risk of conflict between individuals that have never seen each other.

The megaherbivores — terrestrial mammals that feed on plants and weigh over 1000 kg (elephants, hippos

and rhinos) — are a major source of concern. Rhinos remain critically endangered<sup>2</sup>. Elephants are also under pressure from human activities, yet our extensive knowledge of their biology allows conservationists to devise appropriate management measures<sup>3</sup>. The third African megaherbivore, the hippopotamus, is not yet listed as endangered, but its populations have declined dramatically in recent decades<sup>4</sup>. This amphibious animal shares its life between land and water and has a unique role in the ecosystem mainly because of its impact on the flux of energy and matter between the two environments<sup>5</sup>. The biology of the hippopotamus is still mysterious in many respects, and population management methods remain largely empirical<sup>6</sup>.

Studying the behavioral biology of hippos in the wild is complicated. It is difficult, if not impossible, to identify and mark individuals and sometimes highly challenging to locate them. Hippos feed on land mainly at night and are rather solitary. During the day they



**Figure 1. Playback experiments on hippos.**

(A) Map of the Maputo Special Reserve (Mozambique, Austral Africa) with location of the tested hippo groups. (B) Examples of hippo vocalizations used for the playback experiments (three different individuals). (C) Top: typical hippo group. Bottom: approach toward the loudspeaker and marking (dung spraying followed by threatening display). (D) Behavioral reaction of hippos to conspecific vocalizations (familiar: call from an individual of the tested group; neighbor: call from an individual of a group from the same lake; stranger: call from an individual of a distant group). Left: Response strength represents an integrative measure of the hippo reaction to calls, calculated using a principal component analysis from all behavioral variables. Right: Diagram reporting the behavioral scores for each of the three considered behaviors (approach toward the loudspeaker, number of vocalizations, and marking behavior by defecation; see Supplemental information for details).





gather in groups in the water. Hippo groups are socially structured around a dominant male, a variable number of females and their young, and some peripheral males<sup>7</sup>. However, it is unclear whether individuals in a pod form a stable group defending a territory or pods are organized in a fission-fusion manner with individuals moving from one pod to another. In any case, the hippo social system appears to rely on communication signals — hippos are very vocal — whose role and meaning remain almost unexplored<sup>8</sup>.

Here, we focused on the most common hippo vocalization, the ‘wheeze honk’, a loud call heard over long distances that is assumed to be important for social cohesion and communication between groups, but whose actual function remains unknown (Supplemental information). Based on the assumption that hippo groups are territorial entities, defended by the dominant male but also potentially by other individuals, including females, we tested the hypothesis that the wheeze honk could signal the identity of the sender and thus enable behavioral decisions by the receiver individuals.

We worked in the *Maputo Special Reserve* (Mozambique; Figure 1A). For each group of hippos (minimum number of individuals = 3; maximum = 22), we first recorded spontaneous vocalizations and then conducted playback experiments (Supplemental information).

We conducted three types of playback test on groups of hippos: one with a call from the group, another with a call from a different group present on the same lake, and another with a call from a distant stranger group (Figure 1B). Of the seven groups tested, five received all three stimuli. Two groups received only the familiar and stranger stimuli (one group had no neighbors in its lake and one has not been retested due to experimental constraints; Supplemental information). The order of the tests was balanced among the groups. The signals were played from the shore, around 70–90 meters away from the group (mean duration of a playback session = 36 minutes, min–max = 15–75 min).

Hippos respond to played back call (by calling back, approaching and/or marking by defecation), but their response depends on the category of the stimulus (Figure 1C,D; Supplemental information).

The overall intensity of the behavioral response is lowest in response to a call from an individual of the same group and highest in response to a call from an individual belonging to a stranger group (linear mixed model, Wald  $X^2 = 17.55$ ,  $p < 0.001$ ; see Supplemental information for detailed statistics). The nature of the response also changes between stimuli. Whereas individuals responded to calls from any group, marking behavior (dung spraying) is modulated by the category of the calls (cumulative mixed model: Wald  $X^2 = 11.47$ ,  $p = 0.003$ ). Stranger group calls induce more marking than calls from an individual of the same group (multiple comparisons tests:  $Z = 2.41$ ,  $p = 0.042$ ), while there is no significant difference between reactions to the calls from the same group or from a neighboring group (multiple comparisons tests:  $Z = 0.40$ ,  $p = 0.915$ ).

Individuals in territorial animals often react less aggressively to a known individual from a neighboring territory than to a stranger (‘dear-enemy effect’). Sometimes, the opposite is observed and it is the neighbor that is more strongly repelled (‘nasty neighbor’). Our experiments suggest that in hippos, the arrival of a stranger individual is perceived as more threatening than that of a neighbor.

Complementary experiments to assess sound level and sound propagation revealed that the wheeze honk can propagate more than 1 km away (Figure S1). Although the acoustic features carrying the vocal signature may be altered during long distance transmission, it is thus likely that these animals can learn and recognize the voices of neighboring individuals living on the same lake.

In short, we showed that hippos use vocal recognition to manage their inter-group relationships, a strategy already observed in other large mammals where competition is intense and that limits physical fights with often irreparable consequences<sup>9</sup>. While relocating endangered animals to maintain populations above critical levels is increasingly common<sup>10</sup>, our results suggest that precautions should be taken during such relocations with hippos. Before transferring a group of hippos to a new location, a potential precaution could be to broadcast their voices from a loudspeaker at a distance from the groups already present so that

they get used to them and their level of aggressiveness gradually decreases. Getting the animals to be relocated accustomed to the voices of their new neighbors could also be considered.

#### SUPPLEMENTAL INFORMATION

Supplemental information includes one figure, one table, experimental procedures and discussion, and can be found with this article online at <https://doi.org/10.1016/j.cub.2021.12.017>.

#### REFERENCES

- Enquist, B.J., Abraham, A.J., Harfoot, M.B.J., Malhi, Y., and Doughty, C.E. (2020). The megabiota are disproportionately important for biosphere functioning. *Nat. Commun.* *11*, 699.
- Gross, M. (2018). Last call to save the rhinos. *Curr. Biol.* *28*, R1–R3.
- Wall, J., Wittemyer, G., Klinkenberg, B., LeMay, V., Blake, S., Strindberg, S., Henley, M., Vollrath, F., Maisels, F., Ferwerda, J., et al. (2021). Human footprint and protected areas shape elephant range across Africa. *Curr. Biol.* *31*, 2437–2445.
- Lewis, R., and Pluháček, J. (2017). *Hippopotamus amphibius*. The IUCN Red List of Threatened Species 2017: e.T10103A18567364. <https://dx.doi.org/10.2305/IUCN.UK.2017-2.RLTS.T10103A18567364.en>. Downloaded on 4 July 2021.
- Schoelynck, J., Subalusky, A.L., Struyf, E., Dutton, C.L., Unzué-Belmonte, D., Van de Vijver, B., Post, D.M., Rosi, E.J., Meire, P., and Frings, P. (2019). Hippos (*Hippopotamus amphibius*): The animal silicon pump. *Sci. Adv.* *5*, eaav0395.
- Utete, B. (2020). A review of some aspects of the ecology, population trends, threats and conservation strategies for the common hippopotamus, *Hippopotamus amphibius* L, in Zimbabwe. *Afric. Zool.* *55*, 187–200.
- Klingel, H. (2013). *Hippopotamus amphibius* common hippopotamus. In *Mammals of Africa Volume VI: Hippopotamuses, Pigs, Deer, Giraffe and Bovids*, J. Kingdon and M. Hoffmann, eds. (London: Bloomsbury Publishing), pp. 68–77.
- Maust-Mohl, M., Soltis, J., and Reiss, D. (2015). Acoustic and behavioral repertoires of the hippopotamus (*Hippopotamus amphibius*). *J. Acoust. Soc. Am.* *138*, 545–554.
- Mathevon, N., Casey, C., Reichmuth, C., and Charrier, I. (2017). Northern elephant seals memorize the rhythm and timbre of their rivals’ voices. *Curr. Biol.* *27*, 2352–2356.
- Berger-Tal, O., Blumstein, D.T., and Swaisgood, R.R. (2020). Conservation translocations: a review of common difficulties and promising directions. *Anim. Conserv.* *23*, 121–131.

<sup>1</sup>Equipe de Neuro-Ethologie Sensorielle ENES/CRNL, University of Saint-Etienne, CNRS, Inserm, Saint-Etienne, France. <sup>2</sup>Equipe Cognition Auditive et Psychoacoustique CAP/CRNL, University Lyon 1, CNRS, Inserm, Lyon, France. <sup>3</sup>Departamento de Biologia Animal and cE3c — Centre for Ecology, Evolution and Environmental Changes, Faculdade de Ciências, Universidade de Lisboa, Lisbon, Portugal. <sup>4</sup>Institut Universitaire de France, Paris, France. <sup>5</sup>Co-senior authors. <sup>6</sup>Lead contact.

\*E-mail: [pjfonseca@fc.ul.pt](mailto:pjfonseca@fc.ul.pt) (P.F.); [mathevon@univ-st-etienne.fr](mailto:mathevon@univ-st-etienne.fr) (N.M.)

---

# Auditory perception and decision-making in crocodilians

---

**Keywords :** *Crocodilians, Bioacoustics, Auditory perception, Sound detection, Sound categorization, Interspecific communication, Decision making*

Crocodilians face a flow of auditory information, which they must detect and identify. To understand how they analyze their auditory environment, I conducted a series of studies to answer three questions: 1) how do crocodilians detect a sound signal in background noise; 2) what are their abilities to form categories of sounds along a sound continuum; 3) what are the acoustic bases of the recognition of a biologically relevant sound signal: the distress call of juveniles? I first conducted experiments on acoustic signal detection in noisy environment. My experiments show that the spatial separation of the signal from the noise source improves their detection performance, highlighting a spatial unmasking mechanism that allows crocodilians to improve their auditory scene analysis. I then turned my attention to information coding and how young crocodilians can identify a call. Whether naive or conditioned, crocodilians faced with an acoustic continuum ranging from a frog call to a crocodile call classify each acoustic variant into one of these two categories based solely on the spectral envelope of the sounds, demonstrating that they are able to use categorization to quickly respond to information in a signal stream. Finally, I show that crocodilians are sensitive to the degree of distress encoded in the vocalizations of phylogenetically distant vertebrates, demonstrating their abilities to take advantage of acoustic information present in their sound environment. All of these studies contribute to better understand how crocodilians perceive their acoustic environment and demonstrate that they can use the information encoded in acoustic signals to guide their decision-making.

---

# Perception auditive et prise de décision chez les crocodiliens

---

**Mots-clés :** *Crocodiliens, Bioacoustique, Perception auditive, Détection des sons, Catégorisation des sons, Communication interspécifique, Prise de décision*

Les crocodiliens font face à un flux d'informations auditives qu'ils doivent détecter et identifier. Afin de comprendre comment ils analysent leur environnement sonore, j'ai mené une série d'études pour répondre à 3 questions : 1) comment les crocodiliens détectent un signal sonore dans un bruit de fond ; 2) quelles sont leurs capacités à former des catégories de sons le long d'un continuum ; 3) quelles sont les bases acoustiques de l'identification d'un signal sonore biologiquement pertinent : le cri de détresse de jeunes. J'ai d'abord mené des expériences de détection de signaux en environnement bruyant, qui montrent que la séparation spatiale du signal de la source de bruit améliore les performances de détection des crocodiles, mettant en évidence un mécanisme de démasquage spatial qui contribue à leur analyse de la scène auditive. Je me suis ensuite intéressée au codage de l'information et à la manière dont les jeunes crocodiliens identifient un cri. Qu'ils soient naïfs ou conditionnés, les crocodiliens confrontés à un continuum acoustique allant d'un cri de grenouille à un cri de crocodile classent chaque variante acoustique dans l'une de ces 2 catégories en se basant uniquement sur leur enveloppe spectrale, démontrant qu'ils sont capables d'utiliser la catégorisation pour réagir rapidement aux informations d'un flux de signaux. Enfin, je montre que les crocodiliens sont sensibles à la détresse encodée dans les vocalisations de vertébrés phylogénétiquement éloignés, démontrant leur capacité à tirer profit des informations acoustiques présentes dans leur environnement. Toutes ces études contribuent à mieux comprendre comment les crocodiliens perçoivent leur environnement acoustique et démontrent qu'ils peuvent utiliser les informations encodées dans des signaux acoustiques pour guider leur prise de décision.