



**HAL**  
open science

# Sur les traces de l'origine de nos conversations : étude comparative des interactions vocales chez les grands singes

Loïc Pournault

► **To cite this version:**

Loïc Pournault. Sur les traces de l'origine de nos conversations : étude comparative des interactions vocales chez les grands singes. Psychologie et comportements. Université de Rennes, 2021. Français. NNT : 2021REN1B002 . tel-03275008

**HAL Id: tel-03275008**

**<https://theses.hal.science/tel-03275008>**

Submitted on 30 Jun 2021

**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 DE RENNES 1

ECOLE DOCTORALE N° 605

*Biologie Santé*

Spécialité : Neurosciences, Ethologie

Par

**Loïc POUGNAULT**

## **Sur les traces de l'origine de nos conversations : étude comparative des règles d'interactions vocales chez les grands singes**

Thèse présentée et soutenue à Rennes, le 22 Janvier 2021

Unité de recherche : UMR 6552

### **Rapporteurs avant soutenance :**

Cécile Garcia Chargée de Recherche,  
UMR 7206

Maya Gratier Professeure,  
EA 3456

### **Composition du Jury :**

Président  
Michael Greenfield Professor,  
University of Kansas  
ENES/CRNL UMR CNRS 5292

Examineurs  
Laurence Henry Maître de Conférence,  
UMR 6552

Rapporteurs  
Cécile Garcia Chargée de Recherche,  
UMR 7206  
Maya Gratier Professeure,  
EA 3456

Directeur de thèse  
Alban Lemasson Professeur, Directeur d'unité,  
UMR 6552

Co-directrice de thèse  
Florence Levréro Maîtresse de Conférence,  
ENES/CRNL UMR CNRS 5292

### **Invité**

Baptiste Mulot Responsable Vétérinaire, Directeur de  
la Recherche,  
ZooParc de Beauval



*« Le silence même se définit par rapport aux mots, comme la pause en musique, reçoit son sens des groupes de notes qui l'entourent. Ce silence est un moment du langage ; se taire, ce n'est pas être muet, c'est refuser de parler, donc parler encore... »*

*Jean-Paul Sartre, Situations II, 1948*



# *Remerciements*

*Ce n'est pas le but qui compte, c'est le chemin.* Ce chemin peut toutefois prendre des tournures insoupçonnées au gré des rencontres. Tant de personnes ont contribué à l'aboutissement de ce projet que je ne sais par où commencer. Ne voyez en aucun cas, dans l'ordre d'apparition, une quelconque préférence pour l'un/l'une ou pour l'autre. Chacun, à sa façon, a apporté sa pierre à l'édifice.

A **Florence Levréro** et à **Alban Lemasson**, mes directeurs de thèse. Un grand merci à vous d'être venus me récupérer au fond de la jungle Péruvienne, de m'avoir fait confiance et de m'avoir soutenu avec optimisme tout au long de ce projet. Je vous remercie également pour les échanges que nous avons pu avoir, qui ont fait grandir et murir ma réflexion. Enfin, je vous suis particulièrement reconnaissant d'avoir toujours su faire preuve de bienveillance à mon égard, dans les bons comme dans les mauvais moments.

A **Baptiste Mulot**. Merci à toi de m'avoir accueilli à bras ouverts au sein du ZooParc de Beauval, d'avoir rendu chaque séjour dans ta structure particulièrement bienveillant et chaleureux. Je tenais aussi à souligner la disponibilité et l'accessibilité dont tu as toujours su faire preuve malgré ton emploi du temps surchargé, merci. Sans toi, cette thèse n'aurait pas le même visage.

A **Cécile Garcia** et **Maya Gratier**. Pour l'honneur qu'elles me font d'être rapporteuses de ce travail, mais aussi à **Laurence Henry** et **Michael Greenfield** pour avoir accepté de prendre part à mon jury.

A **Marie Bourjade** et **Isabelle Charrier**. Pour avoir accepté de faire partie de mon comité de thèse, pour leurs encouragements, leurs conseils et les échanges que nous avons pu avoir lors de ces temps de parole, mais aussi en dehors.

A **Rodolphe Delord**, actuel directeur du ZooParc de Beauval. Qui m'a permis de mener cette thèse au sein de son établissement.

A **Nicolas Mathevon**. Qui m'a accueilli au sein de son laboratoire et pour sa bienveillance. Plus largement, un grand merci à **tous les membres de l'ENES** (Clément, Léo, Thibaut, Lucie, Mathilde, David, Vincent, Joël, Marilyn, Niko et pardonnez-moi si j'oublie quelqu'un) pour leur accueil, nos échanges et les quelques soirées que j'ai pu passer à vos côtés.

A **Caroline Coste, Véronique Guyot, Géraldine Legoff et Antoine L'Azou**. Pour leur organisation, leur disponibilité et leurs conseils qui ont déjoués bon nombres de soucis auxquels j'aurais pu être confrontés.

A **Hélène Cartaud, Catherine Racineux, Daniel Cluzeau et Régis Super** ainsi qu'à l'ensemble des **Stationnaires**. Merci de faire de la Station Biologique un havre de paix, un petit îlot chaleureux.

A **Manon, Agnès, Antoine, Ludivine et Thomas**, les soigneurs de l'équipe Gorilles du ZooParc de Beauval. Merci pour votre accueil, votre soutien et les sacrifices auxquels vous avez dû faire face pour me permettre de mener à bien mes études, qu'elles soient ou non couronnées de succès. Je reste ébahi par votre travail, votre engagement quotidien. Sans vous, rien de tout cela n'aurait existé, merci ! (Agnès, je conserve dans un coin de ma tête l'idée ton idée recherche sur les gorilles)

A **Julie, Sam, Seb et Chléa**, les soigneurs de l'équipe Chimpanzés/Orangs-outans du ZooParc de Beauval. Pour le temps que vous avez su m'accorder, pour la patience que vous avez eu à mon égard et pour le travail quotidien que vous fournissez pour rendre le plus agréable possible la vie des animaux dont vous avez la charge, merci.

To **Jord, Melissa**, the caretakers of the orangs-outans at the Ouwehand Dierenpark of Rhenen. Thank you again for your warm welcome, for all these very tasty « Christmas cakes »! You are doing an extraordinary job.

A **Arnaud Rossard et Phillipe Bec**. Merci d'être là, d'apporter un peu de piment à la vie Paimpontaise, pour votre bonne humeur, votre humour et votre recul sur les choses.

A l'ensemble des **membres du laboratoire EthoS**. Pour leur bonne humeur, les jeux partagés le midi, les rires et les discussions plus ou moins sérieuses que nous avons pu avoir.

A **Elisa Demuru**. Ces quelques jours en ta compagnie et nos nombreux échanges m'ont fait grandir, aussi bien professionnellement que personnellement. Merci d'avoir partagé avec moi le regard que tu portes sur le monde.

A **Melissa Sébilleau** et **Maël Leroux**. Que vais-je bien pouvoir trouver à dire sur vous deux... Votre humour cynique et les galères que nous avons partagées depuis le Master 2 m'ont indéniablement permis de tenir le coup dans les moments difficiles. Malgré la distance qui peut nous séparer depuis quelques années, je ne me suis jamais senti aussi proche de vous !

A **Juliette**, ma « binôme » de thèse. Avant toute chose, j'espère que tu es fière de toi, de tout ce que tu as accompli au cours de ton projet. Tu as su soulever des montagnes. Pendant ces quelques années, tu as été un vrai rayon de soleil pour moi. Ta joie de vivre perpétuelle, ton positivisme acharné et ton humour graveleux ont été un cocktail énergisant et un bol d'air frais à chaque instant. Continue de t'émerveiller face aux grands yeux humides de toute la faune que tu croises, de collectionner les timbres représentant des œufs et de la météo bretonne. Merci à toi, Juliette.

A **Noémie**, le sourire incarné. Tu m'étonneras toujours par le recul que tu arrives à prendre sur les choses, sur la façon dont tu perçois la vie. Continue de croire en la chance que tu as et que mérites amplement. Merci pour tout ce que tu m'as appris.

A **Pablo**, la vitamine. Même si ton passage dans le bureau des doctorants n'aura pas été très long, tu auras su insuffler une énergie dont toi seul a le secret. J'espère avoir la chance de te voir plus souvent à l'océan, au large, pour avoir le plaisir de partager avec toi quelques déferlantes.

Aux doctorants Rennais, **Marion**, **Delphine** et **Christiane**. Merci à vous pour votre enthousiasme et votre soutien. Vous rencontrez au cours de cette thèse aura été une vraie chance pour moi. Courage à vous et aux futurs thésards !

Aux anciennes doctorantes, **Amélie**, **Mathilde** et **Serenella**, qui nous ont passé le flambeau. Merci pour votre écoute, votre aide et vos encouragements. Un merci particulier à Mathilde pour m'avoir légué son bureau près de la fenêtre à travers laquelle je pouvais observer les sittelles et autres passereaux, ce qui n'a pas facilité ma concentration sur ces derniers mois.



A **Julien Paulet** et **Thomas Ferrari**, les stagiaires que j'ai eu la chance de rencontrer et d'encadrer. J'espère ne pas avoir été un encadrant trop exigeant. Merci pour l'aide considérable que vous m'avez apporté au cours de cette thèse.

A **ma petite sœur** et **mes parents**. Pour m'avoir écouté patiemment parler de mes singes, de mes « expé » et de mes « obs » qui devaient leur paraître particulièrement abstraites. Pour leur soutien inconditionnel malgré la distance et pour m'avoir donné l'envie d'être heureux et de profiter de chaque jour. Merci à vous.

Enfin, je tenais à te remercier, **Margaux**, d'avoir été là à chaque instant, partageant mes moments de joie, de doute ou même de tristesse. Tu as été un phare dans la tempête qu'a pu parfois représenter ce projet. Il me sera difficile de trouver les mots pour exprimer toute la chance que j'ai de t'avoir à mes côtés.

Je n'oublie pas l'ensemble des primates non-humains avec lesquels j'ai eu la chance d'interagir pendant ces quelques années. Je ne peux réellement les remercier pour leur participation, car ma présence leur a été imposée. Toutefois, je leur suis très reconnaissant de m'avoir émerveillé à chaque rencontre et de m'avoir toléré auprès d'eux.





# Sommaire

<b>Chapitre I - Introduction générale</b> .....	15
1. La vie sociale : moteur de l'évolution de la communication vocale .....	17
1.1. Quelle définition pour la complexité sociale ?.....	17
1.2. Quelle définition pour la complexité vocale ? .....	19
1.3. L'intérêt d'une approche comparative .....	21
2. La comparaison primates humains et non-humains : le débat sur les origines du langage .....	22
2.1. Des controverses historiques .....	22
2.2. Les origines non-vocales .....	24
2.2.1. Une origine gestuelle .....	24
2.2.2. Une origine labiale.....	25
2.3. Les origines vocales.....	26
2.4. Vers un scénario évolutif multi-modal et multi-causal du langage .....	28
3. La place des études sur les interactions vocales dans le débat sur les origines du langage .....	30
3.1. Les <i>chorus</i> .....	31
3.2. Les duos.....	33
3.3. Les échanges vocaux organisés sur le tour de parole .....	34
Article 1 : Conversation among primate species .....	37
4. Problématique et objectifs de la recherche .....	63
4.1. Problématique.....	63
4.2. Choix des modèles d'études .....	64
4.3. Objectifs de l'étude.....	64
<b>Chapitre II - Méthodologie générale</b> .....	67
1. Les grands singes non-humains .....	69
1.1. Les chimpanzés .....	69
1.1.1. Phylogénèse, répartition géographique et écologie .....	69
1.1.2. Système social et territorialité.....	70
1.1.3. Communication .....	71
1.2. Les bonobos.....	73
1.2.1. Phylogénèse, répartition géographique et écologie .....	73
1.2.2. Système social et territorialité.....	74
1.2.3. Communication .....	75
1.3. Les gorilles .....	76
1.3.1. Phylogénèse, répartition géographique et écologie .....	76
1.3.2. Système social et territorialité.....	77

1.3.3. Communication .....	78
1.4. Les orangs-outans .....	79
1.4.1. Phylogénèse, répartition géographique et écologie .....	79
1.4.2. Système social et territorialité .....	80
1.4.3. Communication .....	81
2. Sujets d'étude et conditions de vie .....	82
2.1. En parcs zoologiques .....	83
2.2. A l'état sauvage .....	85
3. Collecte de données .....	86
3.1. Enregistrements acoustiques et approche observationnelle .....	87
3.1.1. La production vocale .....	87
3.1.2. Facteurs socio-démographiques et contextuels .....	93
3.2. Approche expérimentale .....	94
3.2.1. Paradigme de violation des attentes .....	94
3.2.2. Objectif de l'expérimentation .....	95
3.2.3. Dispositif expérimental .....	95
3.3. Considérations statistiques .....	97
<b>Chapitre III - Pertinence sociale des règles d'interactions vocales .....</b>	<b>99</b>
Article 2 : Breaking conversational rules matters to captive gorillas : a playback experiment .....	101
<b>Chapitre IV - Organisation temporelle des productions vocales et influences sociales .....</b>	<b>117</b>
Article 3 : Preliminary investigation of temporal calling patterns of a captive group of chimpanzees .....	119
<b>Chapitre V - Influence du système social sur l'émergence des échanges vocaux organisés .....</b>	<b>159</b>
Article 4 : "Conversational rules" are driven by social pressures, say studies on great apes .....	161
<b>Chapitre VI - Discussion générale .....</b>	<b>229</b>
1. Principaux résultats .....	231
2. Développement ontogénétique des échanges vocaux .....	233
2.1. Contrôle moteur et développement des structures cérébrales .....	234
2.2. Développement sous influences sociales .....	237
2.2.1. Apprentissage social des règles d'interactions .....	238
2.2.2. Influence de la motivation sociale .....	240
3. Importance des règles et fonctions des échanges vocaux .....	242
3.1. Importance du respect des règles .....	242
3.2. Fonctions des échanges vocaux organisés .....	243
4. Emergence des échanges vocaux .....	245
4.1. Des facteurs sociaux essentiels à la mise en place des échanges vocaux ? .....	245
4.2. L'apparition des échanges vocaux : convergence évolutive ou caractère ancestrale ? .....	248
5. Perspectives majeures .....	250

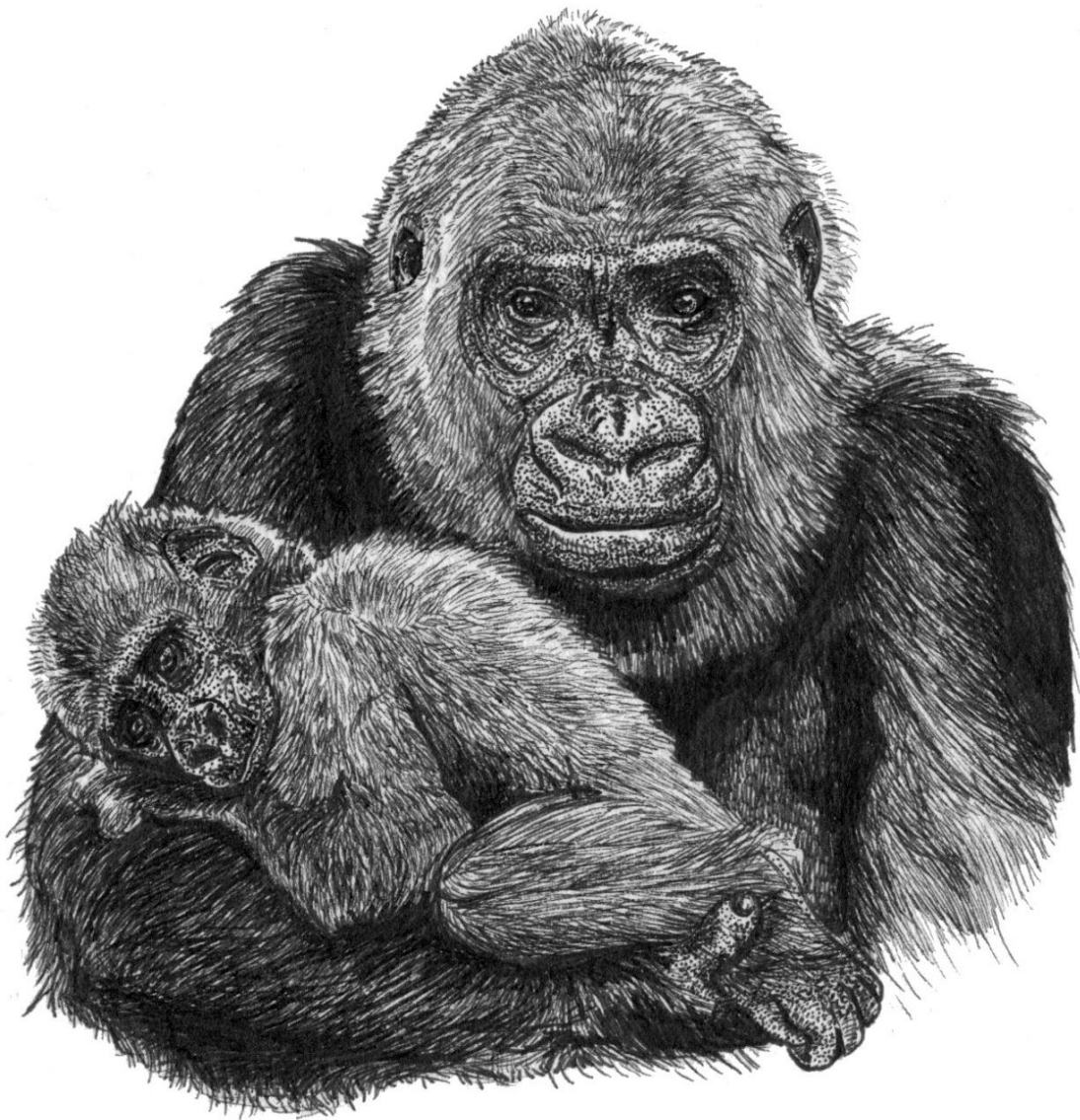
---

5.1. Des précisions à apporter.....	251
5.2. Les échanges vocaux organisés : fruit d'une intentionnalité sous-jacente ou simple contagion émotionnelle ?.....	253
6. Conclusion.....	255
<b>Chapitre VII - Références bibliographiques .....</b>	<b>259</b>
<b>Chapitre VIII - Annexes .....</b>	<b>299</b>
Annexe 1 : Liste des publications.....	301
Annexe 2 : Communications orales .....	303
Annexe 3 : Activités de diffusion de la culture scientifique .....	309
Annexe 4 : Encadrement de projets, Responsabilité collectives & Formations suivies .....	311



# Chapitre I

## Introduction générale



© Illustration de Louise Prévot





## 1. La vie sociale : moteur de l'évolution de la communication vocale

Au cœur des théories sur les origines du langage humain réside la question du rôle joué par la vie sociale (Dunbar 1996). Dès la fin du XIX<sup>ème</sup> siècle, Darwin, dans son ouvrage « *The expression of the emotions in man and animals* », suggérait l'existence d'un lien entre l'évolution de la socialité et de la communication :

« *Naturalists have remarked, I believe with truth, that social animals, from habitually using their vocal organs as a means of intercommunication, use them on other occasions much more freely than other animals* »

Depuis, l'hypothèse d'une co-évolution de la complexité sociale-vocale s'est précisée. Elle suggère que plus les groupes d'individus ont un système social complexe, plus ils requièrent un système de communication complexe afin de réguler la diversité des interactions et des relations inter-individuelles (Freeberg *et al.* 2012).

Néanmoins, l'appréhension de la notion de complexité a posé, et pose encore, de nombreux défis. Bien que le terme de « complexité » soit couramment utilisé pour caractériser le vivant, il est rarement défini par les auteurs (*e.g.* Adami 2002). Toutefois, il semble se dessiner un consensus selon lequel un système variable et hétérogène est considéré comme complexe de par le nombre d'éléments qui le compose et l'inconstance de leur arrangement (*e.g.* McShea 1991).

### 1.1. Quelle définition pour la complexité sociale ?

Un système social peut se définir en trois dimensions (Kappeler et van Schaik 2002) :

- (i) La structure démographique du groupe, qui correspond à la taille du groupe, au nombre individus de différents âges et sexes et au niveau de cohésion spatiale des individus au cours du temps.
- (ii) L'organisation sociale du groupe, qui peut être définie selon trois niveaux (Hinde 1976) que sont les interactions sociales (*i.e.* toutes actions à la valence négative ou positive d'un individu vers un autre), les relations sociales (*i.e.* somme des interactions qui donne sa nature de la relation), et le réseau social (*i.e.* le réseau de relations existant entre les membres d'un groupe).
- (iii) Le système de reproduction, pouvant être monogame (*i.e.* un mâle et une femelle qui s'accouplent uniquement en période de reproduction : Martinez et Garcia 2020)

ou polygame, incluant la polyandrie (*i.e.* une femelle s'accouple avec plusieurs mâles pendant la période de reproduction : Martinez et Garcia 2020) et la polygynie (*i.e.* un mâle s'accouple avec plusieurs femelles pendant la période de reproduction : Martinez et Garcia 2020).

A partir de ces différentes caractéristiques, certains auteurs ont tenté de comparer les systèmes sociaux par une approche qualitative. Pour certains, une société soumise à une dynamique de fission-fusion (*i.e.* formation de sous-groupes temporaires dont la taille et la composition sont labiles) peut être considérée comme plus complexe que les sociétés qui ne sont pas soumises à cette dynamique (Aureli *et al.* 2008). De même, un niveau de complexité supérieure est alloué aux groupes stables par rapport aux groupes instables, car la stabilité favoriserait l'émergence de comportements coopératifs (Wilkinson 2003). La complexité attribuée aux différents systèmes de reproduction anime également les débats. Pour certains auteurs, les groupes multi-mâles / multi-femelles dominent par leur complexité les groupes monogames (Shultz et Dunbar 2006), alors que d'autres chercheurs avancent l'idée selon laquelle le lien de long terme unissant les couples monogames est propice à la coordination des individus, ce qui témoignerait d'une complexité supérieure (Freeberg *et al.* 2012; Peckre *et al.* 2019). En outre, la complexité à accorder aux différents types de hiérarchies (*i.e.* linéaire et non-linéaire) ne fait pas non plus l'objet d'un consensus (*e.g.* MacLean *et al.* 2008; Peckre *et al.* 2019; Taborsky et Oliveira 2012).

Au regard de la difficulté que semble représenter l'attribution d'un niveau de complexité en fonction de critères qualitatifs, une approche quantitative est peut-être plus objective. Le nombre d'individus composant le groupe social est souvent utilisé comme critère de complexité sociale (Freeberg *et al.* 2012; Lehmann et Dunbar 2009). Ainsi, plus un groupe comporte d'individus, plus ces derniers vont interagir entre eux dans des contextes variés. Toutefois, ce critère est aisément critiquable. Par exemple, il est difficile d'attribuer une complexité supérieure à un groupement grégaire de poissons qui ne semblent pas exprimer de préférences particulières pour l'un ou l'autre de leurs congénères (Parrish *et al.* 2002), à l'inverse d'un couple de primates (Maninger *et al.* 2017). Pour pallier à cela, des auteurs proposent de compléter ce critère (*i.e.* effectif du groupe) par le nombre de rôles sociaux (Blumstein et Armitage 1997; Pollard et Blumstein 2012), comme par exemple les individus venant en aide à la mère allaitante. Pourtant, malgré l'ajout de ce second critère, le modèle reste critiquable car il ne tient pas compte des interactions ayant lieu au sein du groupe. Freeberg et collaborateurs (2012) définissent alors un système social complexe comme un système dans lequel les mêmes

individus interagissent dans des contextes divers et variés, avec de multiples individus, et de façon répétée avec les mêmes individus au cours du temps. Pour réduire encore l'imprécision de cette définition, Bergman et Beehner (2015) en suggèrent une plus épurée se réduisant au nombre de relations différentes qu'entretient chaque individu. En ce sens, Fischer et collaborateurs (2017a) ont récemment proposé une nouvelle méthode qui tient compte de quatre indices sociaux (*i.e.* diversité des patterns comportementaux, indice composite dyadique de socialité, fréquence relative d'interaction et la valence de ces interactions) pour identifier quatre types de relations sociales au niveau dyadique (*i.e.* rares interactions agonistiques, rares interactions affiliatives, interactions ambivalentes moyennement fréquentes et fréquentes interactions affiliatives) et calculer un indice de la diversité relationnelle au niveau individuel.

Comme nous venons de le voir, les interactions sociales régissent la vie sociale. Cela implique que les individus doivent être en mesure de reconnaître et de se rappeler de leurs interactions passées, de répondre de façon pertinente aux interactions ayant eu lieu entre des membres tiers du groupe (*e.g.* l'individu réagira différemment face à un congénère récemment agressé par rapport à un congénère non-agressé), ou encore d'évaluer le comportement des autres membres du groupe et de réagir en conséquence (Freeberg *et al.* 2012). Pour cela, la capacité à communiquer apparaît comme cruciale (Hauser 1996; Owings et Morton 1998) et semblerait gagner en complexité parallèlement à la complexification de la socialité des espèces.

## 1.2. Quelle définition pour la complexité vocale ?

La communication peut être considérée comme un acte social qui permet de transmettre de l'information, sous la forme d'un signal, d'un émetteur vers un receveur (Liebal *et al.* 2012; Smith 1965; Wiley 2012). Le signal est considéré comme le support physique de l'information (Bradburry et Vehrencamp 1998), et il peut transiter par différents canaux : chimiques, électriques, visuels, mécaniques ou encore acoustiques (Bradburry et Vehrencamp 1998). Le signal est alors perçu par le receveur qui va décoder l'information, pouvant induire une réponse de sa part (Bradburry et Vehrencamp 1998). C'est la réponse comportementale du receveur qui nous permet d'inférer la signification du signal transmis (Bradburry et Vehrencamp 1998). Il est toutefois important de noter que la signification d'un signal ne dépend pas uniquement de sa structure (*i.e.* encodant l'information), mais aussi du contexte dans lequel le signal est produit (Smith 1965).

Dans le cas des sons, le signal se définit selon plusieurs paramètres acoustiques dont les principaux sont la fréquence (*i.e.* la hauteur du son, en Hertz), l'amplitude (*i.e.* intensité du son, en décibels) et la durée (*i.e.* temps durant lequel un signal sonore est produit, généralement en

millisecondes). Les différents signaux sonores qu'une espèce peut produire constituent son répertoire acoustique. Dans l'étude des vocalisations, ce répertoire vocal peut être considéré comme gradé ou discret (Hammerschmidt et Fischer 1998; Wadewitz *et al.* 2015). Le terme de répertoire discret est employé quand il existe une distinction claire entre les différents types de vocalisations (Hammerschmidt et Fischer 1998; Wadewitz *et al.* 2015). A l'inverse, ce répertoire est qualifié de gradé s'il n'existe pas de frontières marquées entre les différents types de vocalisations, mais plutôt un *continuum* entre les différentes structures acoustiques (Hammerschmidt et Fischer 1998; Wadewitz *et al.* 2015).

Traditionnellement, la complexité d'un système communicatif se mesure par le biais du nombre d'éléments structurellement et fonctionnellement différents (*i.e.* grande taille du répertoire communicatif), ou par la quantité d'informations qu'il permet de transmettre (Freeberg *et al.* 2012). Ainsi, plusieurs méthodes non-exclusives existent pour appréhender cette complexité de communication.

Concernant la communication vocale, un nombre important de types de vocalisations différentes est souvent considéré comme étant le témoin d'un degré important de complexité vocale (Gustison *et al.* 2012; McComb et Semple 2005). Le décompte du nombre de types de vocalisations distinctes au sein d'un répertoire se réalise principalement par inspection visuelle. Une approche quantitative, par le biais d'algorithmes, se développe également depuis quelques années (Fischer *et al.* 2017b). Ces deux méthodes (*i.e.* inspection visuelle et algorithmes) restent toutefois particulièrement subjectives. De même, la frontière floue entre les signaux vocaux dans les répertoires gradés rend la tâche encore plus complexe. Bien que Fischer et collaborateurs (2017b) aient suggéré une nouvelle méthodologie permettant d'évaluer le degré de gradation entre deux vocalisations, il n'existe pas, à l'heure actuelle, de mesures quantitatives globales de la gradation du répertoire vocal (Peckre *et al.* 2019).

Le contenu informatif des signaux de communication produits par les individus peut être une seconde piste de réflexion pour évaluer la complexité d'un système de communication, comme le suggère la théorie de l'information qui se base sur l'entropie de Shannon (1948) (aussi nommée théorie mathématique de la communication : Freeberg et Lucas 2012; Shannon et Weaver 1949). Ainsi, plus un système de communication permet de transmettre un grand nombre d'informations, plus il est considéré comme complexe. L'entropie de Shannon (1948) est utilisée par certains auteurs pour quantifier la probabilité d'utilisation d'une vocalisation particulière du répertoire (*e.g.* Bouchet *et al.* 2013; Freeberg 2006). Par exemple, une espèce qui possède un répertoire de dix vocalisations distinctes mais qui en utilise principalement deux

sera considérée comme ayant un degré de complexité vocale inférieure à une espèce qui possède un répertoire équivalent mais qui utilise régulièrement l'ensemble de vocalisations. En revanche, pour que cette théorie de l'information soit valide, les individus doivent posséder les capacités cognitives nécessaires pour percevoir ces différentes vocalisations (Freeberg *et al.* 2012; Manser *et al.* 2014). Pour tirer des conclusions quant à la complexité d'un système de communication par cette approche, il est donc essentiel de tester systématiquement la signification de la variation des signaux pour le récepteur (Manser *et al.* 2014).

En plus des deux méthodes que nous avons vu (*i.e.* nombre d'éléments du répertoire vocal et contenu informatif), des variables complémentaires sont proposées par certains auteurs pour mesurer la complexité communicative d'un système. L'utilisation de différentes modalités de communication peut être une indication de la complexité communicative d'une espèce (Liebal *et al.* 2012). De même, la capacité à encoder des informations sur l'émetteur est également considérée comme un marqueur notable de complexité vocale (Bouchet *et al.* 2013). En outre, certains auteurs ont souligné l'importance de l'aspect référentiel ainsi que les usages contextuels ou intentionnels (Pika 2017). Enfin, nous pouvons supposer que l'organisation temporelle et sociale des interactions vocales pourrait être un élément intéressant à investiguer. Mais, à notre connaissance, personne n'a encore utilisé cette méthode pour évaluer la complexité d'un système vocal.

### 1.3. L'intérêt d'une approche comparative

Les approches comparatives sont indispensables afin d'identifier les pressions de sélection qui peuvent influencer la communication. Les environnements biotiques (*i.e.* social ou non) et abiotiques au sein desquels évoluent les individus peuvent impacter la complexité communicative observée (Freeberg *et al.* 2012; Manser *et al.* 2014). C'est notamment le cas du bruit ambiant enregistré dans un environnement donné, qui peut masquer le signal acoustique produit (Brumm et Slabbekoorn 2005). Ce bruit ambiant peut induire une modulation dans la fréquence de production des sons (*e.g.* Kirschel *et al.* 2009; Slabbekoorn et Peet 2003), ou l'apparition d'un élément sonore au début de signal permettant de capter l'attention de l'audience (*e.g.* Ord et Stamps 2008). La pression de prédation subie par une espèce peut, elle aussi, promouvoir la complexité communicative. Par exemple, cela peut expliquer la mise en place de combinaisons vocales en vue de préciser le danger (*e.g.* mone de Campbell *Cercopithecus campbelli* : Ouattara *et al.* 2009a ; cercopithèque diane *Cercopithecus diana* :

Stephan et Zuberbühler 2008) et son urgence chez certaines espèces de mammifères (écureuil de terre du Cap *Xerus inauris* et suricate *Suricata suricatta* : Furrer et Manser 2009).

Ainsi, l'adoption d'une approche visant à comparer des espèces étroitement apparentées et aux besoins écologiques similaires mais avec des systèmes sociaux différents est importante pour démêler les rôles relatifs de l'environnement et du système social sur la complexité de la communication (Freeberg *et al.* 2012; Manser *et al.* 2014). En suivant cette réflexion, des preuves d'une complexité vocale socialement déterminée ont pu être révélées chez une grande diversité de taxons. Des corrélations ont alors pu être établies entre la taille du répertoire vocal, ou la quantité d'informations contenue dans le répertoire, et la taille du groupe social (*e.g.* chauve-souris : Wilkinson 2003 ; primate : Gustison *et al.* 2012; McComb et Semple 2005), le nombre de catégories démographiques (*e.g.* Sciuridés : Blumstein et Armitage 1997; Pollard et Blumstein 2012), ou encore le système de reproduction (*e.g.* oiseaux : Kroodsmas 1977).

## 2. La comparaison primates humains et non-humains : le débat sur les origines du langage

L'aspect essentiellement social de la communication a fait l'objet d'un consensus depuis plusieurs dizaines d'années (*e.g.* Smith 1969). A l'inverse, la question de l'appartenance exclusive du langage à l'espèce humaine (*Homo sapiens*) fait toujours débat, et ce depuis l'Antiquité.

### 2.1. Des controverses historiques

Dans sa *Politique*, rédigée en -340 avant J-C., Aristote considérait déjà l'humain comme un « animal politique parlant » pour lequel la voix est signifiante. La philosophie s'est par la suite longtemps accaparée la question. C'est notamment le cas de Descartes qui, dans son *Discours de la méthode* (1637), avance l'idée que les animaux non-humains sont capables de « proférer des paroles », mais sans pour autant « parler, c'est-à-dire témoignant de ce qu'ils disent ». Dans le courant de l'année 1871, Darwin suggère, dans son livre « *The descent of man, and selection in relation to sex* », que « l'humain n'est pas le seul animal qui puisse user d'un langage pour exprimer ce qui se passe dans son esprit, et comprendre plus ou moins ce que dit un autre ». Il s'appuie notamment sur l'exemple des primates paraguayens qui, selon lui : « murmurent un ensemble de six mots distincts quand ils sont en état d'excitation, tout en provoquant chez leurs congénères des émotions similaires » (Bouchet *et al.* 2016).

Le langage, ce système communicatif propre à l'humain, tend à se caractériser par rapport aux propriétés qui le distinguent des autres systèmes de communications non-humains

(e.g. Hockett et Hockett 1960; Wacewicz et Żywiczyński 2015). Ainsi, le langage peut être appréhendé comme une entité complexe relevant de capacités telles que la récursivité, la générativité, la symbolique ou encore le déplacement : il apparaît alors comme le propre de l'humain à la vue des découvertes actuelles (e.g. Bickerton 1995; Chomsky 1981; Pinker 1994).

Cependant, en décomposant le langage humain en une somme de propriétés distinctes mais complémentaires, à la fois socio-cognitives (e.g. intentionnalité, flexibilité) et structurelles (e.g. syntaxe, sémantique), il devient possible d'établir des parallèles avec les systèmes de communication des animaux non-humains qui peuvent être informatifs quant aux pressions de sélection ayant influencées l'émergence de ces propriétés et leur diversité (e.g. Fitch 2000; Locke et Snow 1997; Masataka 2003; Meguerditchian *et al.* 2011a; Seyfarth *et al.* 2005; Snowdon et Hausberger 1997). En effet, de par la nature hautement sociale du langage (Lieberman 2007), les recherches menées sur les pré-humains fossiles ne peuvent apporter que peu d'informations quant à son origine. C'est grâce à ces études comparatives que la question des origines du langage a pris un essor considérable sur la scène internationale ces dernières années. Bon nombre des données recueillies sur les propriétés langagières chez des espèces non-humaines proviennent des études sur les oiseaux et les cétacés chez lesquelles un apprentissage social semble nécessaire pour l'acquisition d'un répertoire vocal mature (e.g. Oscine : Baptista et Gaunt 1997 ; orque *Orcinus orca* : Foote *et al.* 2006 ; grand dauphin *Tursiops truncatus* : McCowan et Reiss 1997). Une certaine flexibilité vocale est également reportée chez ces espèces (e.g. étourneau *Sturnus vulgaris* : Hausberger *et al.* 1995 ; cassican flûteur *Gymnorhina tibicen* : Farabaugh *et al.* 1988 ; grand dauphin Janik 2000), tout comme des phénomènes de proto-sémantique faisant référence à l'identité du signaleur (e.g. hirondelle à front blanc *Hirundo pyrrhonota* : Stoddard et Beecher 1983), à des éléments de l'environnement extérieur (e.g. séricorne à courcils blanc *Sericornis frontalis* : Platzen et Magrath 2005 ; poule domestique *Gallus gallus* : Evans et Evans 1999), ou encore des capacités proto-syntaxiques (e.g. mésange de Chine *Parus minor* et cratéropes bicolores *Turdoides bicolor* : Engesser *et al.* 2016; Suzuki *et al.* 2016). Pourtant, la distance phylogénétique séparant les cétacés et les oiseaux du genre *Homo* (Kumar *et al.* 2017) est telle que les similitudes mises en évidence entre le langage humain et leurs systèmes de communication résultent probablement d'une convergence évolutive plutôt que d'un héritage commun (Chow *et al.* 2015). L'étude de ces propriétés langagières chez des espèces distantes est néanmoins un outil non-négligeable pour cibler les contraintes sociales et environnementales ayant favorisées leur émergence (Henry *et al.* 2015).



Les primates non-humains représentent, quant à eux, un intérêt comparatif tout particulier pour l'étude de l'origine du langage de par leur proximité phylogénétique avec l'humain. Les éthologues, psychologues et linguistes sont ainsi parvenus à mettre en évidence des parallèles remarquables entre les capacités communicatives des primates non-humains et plusieurs propriétés langagières, retrouvées de manière universelle dans toutes les cultures humaines (*e.g.* syntaxe : Ouattara *et al.* 2009a; intentionnalité : Schel *et al.* 2013; sémantique: Schlenker *et al.* 2017; flexibilité : Tomasello *et al.* 1994) .

Certaines des propriétés qui composent le langage trouvent une partie de leurs équivalents dans le système de communication non-vocale des primates non-humains, alors que d'autres se dessinent dans leur système de communication vocale. Ainsi, les théories concernant l'origine du langage ont historiquement suivi deux courants distincts opposant les défenseurs d'une origine non-vocale, et principalement gestuelle, face à ceux prônant une origine vocale du langage.

## 2.2. Les origines non-vocales

### 2.2.1. Une origine gestuelle

La théorie d'une origine gestuelle du langage repose sur l'idée que le langage humain trouve ses précurseurs dans l'utilisation de gestes communicatifs produits volontairement chez les primates non-humains (*e.g.* Arbib *et al.* 2008; Corballis 2010; Hewes 1973; Meguerditchian et Vaclair 2008, 2014). L'une des premières caractéristiques de la communication gestuelle des primates non-humains est la grande flexibilité d'apprentissage et d'utilisation des signaux gestuels (Hewes 1973). Bon nombre de projets visant à apprendre le langage oral à des grands singes se sont soldés par des échecs. En revanche les essais d'apprentissages du langage des signes se sont révélés plus concluants (Fitch 2000; Gardner et Gardner 1969; Patterson 1978), soulignant que de nouveaux signaux gestuels communicatifs peuvent apparaître et venir enrichir le répertoire d'un individu adulte (*e.g.* Arbib *et al.* 2008; Corballis 2002; Tomasello 2008). De plus, les adultes expriment un répertoire de gestes communicatifs plus efficace que les juvéniles (*e.g.* Hobaiter et Byrne 2011; Pika *et al.* 2005; Tomasello *et al.* 1985, 1994). Cette flexibilité d'utilisation apparaît aussi en fonction du contexte de production : un même geste pouvant être utilisé dans des contextes distincts et un même contexte pouvant susciter l'utilisation de gestes différents (Tomasello *et al.* 1994).

En plus d'une flexibilité indéniable, la communication gestuelle des primates non-humains présente des marqueurs comportementaux (*i.e.* comportement socialement dirigé

incluant un signal produit dans un but précis et prenant fin une fois le comportement de l'individu cible modifié : Townsend *et al.* 2017) soulignant l'aspect intentionnel de l'utilisation de plusieurs gestes communicatifs (*e.g.* Byrne *et al.* 2017; Leavens *et al.* 2005; Liebal *et al.* 2014a). Certains gestes peuvent aussi exprimer des propriétés référentielles. C'est notamment le cas des gestes de pointage (dit déictiques) qui semblent avoir pour finalité d'orienter l'attention du receveur vers un élément de l'environnement (*e.g.* Genty et Zuberbühler 2014; Krause *et al.* 2018), ou encore des gestes représentant l'action à laquelle ils se réfèrent (dits iconiques) comme les gestes d'invitation chez les bonobos (*Pan paniscus* : Genty et Zuberbühler 2014).

L'un des principaux arguments en faveur de cette théorie reste, néanmoins, d'ordre neuro-anatomique. Chez l'espèce humaine, les zones spécifiques au langage ont été retrouvées dans les aires de Broca et Wernicke, localisées sur l'hémisphère cérébrale gauche (Aboitiz et Garcia 1997; Broca 1865; Horwitz *et al.* 2003; Wernicke 1874; Xu *et al.* 2009). Parallèlement, des observations menées en captivité sur certaines espèces de primates non-humains révèlent, à l'échelle populationnelle, une latéralisation à droite pour la production de gestes communicatifs (*e.g.* Meguerditchian *et al.* 2011b; Meunier *et al.* 2013; Prieur *et al.* 2016a,b). En outre, une asymétrie cérébrale révèle un hémisphère gauche plus développé chez les grands singes, et ce au niveau des aires homologues à celles de Broca et Wernicke (Cantalupo et Hopkins 2001). L'utilisation de techniques d'imageries fonctionnelles a permis de mettre en lumière une augmentation de l'activité de ces régions cérébrales lors de la production de gestes communicatifs chez les chimpanzés (*Pan troglodytes* : Tagliatela *et al.* 2008). Enfin, des neurones miroirs, impliqués dans la reconnaissance et la production de signaux gestuels et oro-faciaux, ont été mis en évidence chez les primates non-humains. Ces neurones se situent dans l'aire cérébrale homologue à celle de la production du langage chez l'humain (Fabbri-Destro et Rizzolatti 2008; Nishitani et Hari 2000; Rizzolatti et Arbib 1998).

### 2.2.2. Une origine labiale

La théorie de l'origine labiale repose sur l'hypothèse que le langage humain trouverait ses précurseurs dans la communication oro-faciale des primates, et plus particulièrement dans la capacité de ces derniers à articuler rythmiquement la langue et les lèvres lors de la production de certaines vocalisations et expressions faciales communicatives (*e.g.* Bergman 2013; Ghazanfar *et al.* 2012; Ghazanfar et Takahashi 2014). En effet, dans les langues occidentales telles que le français ou l'anglais, le langage est soumis à un rythme inhérent à la production des voyelles (induisant une ouverture des lèvres) et des consonnes (induisant une fermeture des

lèvres) qui s'étalonne de 2 Hz à 7 Hz (*i.e.* 2 à 7 cycles d'ouverture/fermeture de bouche par seconde : *e.g.* Chandrasekaran *et al.* 2009; Greenberg *et al.* 2003; Ladefoged et Disner 2012; Munhall et Vatikiotis-Bateson 1998; Ohala 1975; Takahashi *et al.* 2016). Il en résulte une production vocale d'une rythmicité de 3 Hz à 8 Hz, qui semble essentielle pour la bonne compréhension du signal par l'interlocuteur (Chandrasekaran *et al.* 2009; Ghazanfar et Takahashi 2014; Greenberg *et al.* 2003). Les vocalisations de plusieurs espèces de primates manifestent une rythmicité similaire à celle du langage humain (*e.g.* gelada *Theropithecus gelada* : Gustison et Bergman 2017 ; orang-outan *Pongo pygmaeus* : Lameira *et al.* 2015; macaque à face rouge *Macaca articoïdes* : Toyoda *et al.* 2017). De même, la langue, les lèvres et l'os hyoïde s'articulent avec une temporalité équivalente lors de la production de certaines expressions faciales communicatives émises en contextes affiliatifs, telles que le *teeth-chattering* (*i.e.* expression faciale au cours de laquelle les lèvres de l'individu sont complètement découvertes et les dents claquent) chez les macaques à face rouge (Toyoda *et al.* 2017), ou le *lipsmacking* (*i.e.* mouvements verticaux de la mâchoire impliquant généralement une séparation des lèvres) chez les géladas (Bergman 2013), les macaques rhésus (*Macaca mulatta* : Ghazanfar *et al.* 2010, 2012) et très récemment mis en évidence chez les chimpanzés (Pereira *et al.* 2020).

La trajectoire ontogénétique suivie par le *lipsmacking* chez les macaques rhésus n'est pas sans rappeler le développement du langage oral chez l'humain (Morrill *et al.* 2012). En effet, les rythmes inhérents au langage humain et au *lipsmacking* des macaques est lent et variable chez les jeunes, puis s'accélère et se stabilise avec l'âge (humain : Dolata *et al.* 2008; Smith et Zelaznik 2004; macaque rhésus : Morrill *et al.* 2012). Par ailleurs, ce comportement est corrélé avec l'activation d'une région cérébrale homologue à l'air de Broca (Shepherd et Freiwald 2018). Ces arguments ontogénétiques et neuro-anatomiques sont des éléments supplémentaires venant supporter cette théorie.

### 2.3. Les origines vocales

Cette théorie défend l'idée que les vocalisations des primates non-humains sont les précurseurs du langage humain (Lemasson 2011; Masataka 2003; Seyfarth 1987; Snowdon 2009; Zuberbühler *et al.* 2009). Certaines propriétés fondamentales au langage (*e.g.* sémantique, syntaxe, flexibilité et intentionnalité) peuvent en effet être décrites sous leurs formes primitives dans la communication vocale des primates non-humains.

Parmi ces propriétés, la capacité à encoder des informations sur son identité (*e.g.* Fischer *et al.* 2004; Levréro et Mathevon 2013; Noë et Laporte 2014), son état émotionnel (Lemasson *et al.* 2012) et sur l'environnement extérieur (le type d'items alimentaires découverts : *e.g.* Slocombe et Zuberbühler, 2005 ; dans une moindre mesure, le type de prédateurs : *e.g.* Seyfarth et Cheney, 2003) souligne l'aspect référentiel des vocalisations des primates non-humains et permet de proposer des parallèles avec la sémantique. De plus, la propension à combiner les sons de manière non-aléatoire dans le but d'affiner l'information portée par la séquence vocale a été mise en lumière chez certaines espèces (*e.g.* Arnold et Zuberbühler 2006; Clarke *et al.* 2006; Ouattara *et al.* 2009a; Schel *et al.* 2010) et reflète des propriétés syntaxiques et prosodiques (Coye *et al.* 2017).

Malgré un répertoire vocal qui semble fortement déterminé génétiquement (*e.g.* Hammerschmidt et Fischer 2008), il existe des preuves témoignant des capacités de flexibilité sous-jacentes aux vocalisations des primates non-humains. Des variations subtiles de la structure acoustique des vocalisations peuvent se manifester par un phénomène de convergence vocale entre certains membres du groupe (*e.g.* Levréro *et al.* 2015) en fonction de l'âge des individus (*e.g.* Levréro *et al.* 2019), de leurs affinités sociales (*e.g.* De Marco *et al.* 2019; Lemasson et Hausberger 2011; Snowdon et Elowson 1999), du statut social (*e.g.* De Marco *et al.* 2019; Lemasson *et al.* 2016), ou encore induit par une divergence vocale entre groupes voisins (*e.g.* Crockford *et al.* 2004). Des cas d'innovations vocales spontanées ont également été révélés chez certaines espèces (*e.g.* Hopkins *et al.* 2007; Wich *et al.* 2009a). La composition de l'audience semble également influencer les caractéristiques acoustiques des vocalisations (*e.g.* Fedurek *et al.* 2015; Gouzoules *et al.* 1984; Slocombe et Zuberbühler 2007). Par ailleurs, l'utilisation d'une vocalisation spécifique à un contexte précis semble s'acquérir au cours de l'ontogénie (*e.g.* Bouchet *et al.* 2017; Seyfarth et Cheney 1997), et dans de rares cas cette flexibilité développementale peut se manifester dans la structure même des vocalisations (*e.g.* Koda *et al.* 2013; Pistorio *et al.* 2006; Snowdon et Elowson 2001). Ces divers exemples sont autant d'éléments suggérant un apprentissage social et une certaine flexibilité quant à l'utilisation des signaux vocaux chez les primates non-humains. Cette flexibilité reste néanmoins limitée en comparaison à celle s'exprimant dans la communication gestuelle, mais les défenseurs d'une origine vocale du langage estiment pourtant que la flexibilité des vocalisations des primates non-humains est encore fortement sous-estimée (*e.g.* Lemasson *et al.* 2013b; Snowdon 2009). Ces modifications du comportement vocal peuvent amener à questionner la part d'intentionnalité sous-jacente à la communication vocale des primates non-humain.

Les premières études sur ce sujet ont récemment vu le jour. Ainsi, des études menées sur les vocalisations d’alarme des langurs de Thomas (*Presbytes thomasi* : Wich et De Vries 2006) et des cercopithèques à diadème (*Cercopithecus mitis* : Papworth *et al.* 2008) parviennent à démontrer que les individus tiennent compte du niveau de connaissance de l’audience par rapport au danger que représente un prédateur. La première démonstration expérimentale de l’intentionnalité d’un signal vocal a été effectuée chez les chimpanzés où il a été mis en évidence que les vocalisations d’alarmes étaient produites dans un but précis et dirigés socialement (Schel *et al.* 2013). Par la suite, plusieurs études sont venues confirmer ces avancées dans le domaine de l’intentionnalité vocale chez les chimpanzés, couplant les approches observationnelles (*e.g.* Gruber et Zuberbühler 2013) et expérimentales (*e.g.* Crockford *et al.* 2012, 2015).

Enfin, le langage humain est reconnu comme étant par nature un acte hautement social et interactionnel (Levinson 2006). Si bien que certains auteurs, et particulièrement Dunbar (1993, 1996, 2003), suggèrent qu’il a pour fonction première d’établir, de maintenir et de renforcer les liens sociaux affiliatifs entre les individus. Par le biais d’études comparatives menées sur les interactions vocales des primates non-humains, des parallèles avec les conversations humaines ont pu être mis en lumière. En effet, le respect de principes temporeux (*e.g.* alternance des vocalisations des interlocuteurs, évitement du recouvrement des vocalisations) et sociaux (*e.g.* choix non-aléatoire de l’interlocuteur) permettent de qualifier certaines interactions vocales de « conversationnelles » (voir 3.2).

#### 2.4. Vers un scénario évolutif multi-modal et multi-causal du langage

Si la théorie d’une origine vocale et celle d’une origine gestuelle du langage humain prédominent et sont historiquement opposées, de plus en plus de chercheurs ne les considèrent pas comme mutuellement exclusives (*e.g.* Arbib *et al.* 2008; Lemasson 2011; Masataka 2008; Meguerditchian *et al.* 2011a; Meguerditchian et Vauclair 2014; Slocombe *et al.* 2011; Tagliatalata *et al.* 2011), et préfèrent orienter leur réflexion vers la théorie d’une origine multi-modale et multi-causale du langage (*e.g.* Fröhlich *et al.* 2019; Kendon 2017; Liebal *et al.* 2013; Prieur *et al.* 2020; Slocombe *et al.* 2011). Cette théorie alternative et moderne suggère que les propriétés langagières (*e.g.* sémantique, syntaxe, flexibilité et intentionnalité) ont émergé suite à une co-évolution des systèmes de communication gestuel et vocal (*e.g.* Fröhlich *et al.* 2019; Levinson et Holler 2014).

La théorie d’une origine multi-modale du langage s’appuie en partie sur le fait que les primates, humains comme non-humains, intègrent simultanément des signaux de différentes

modalités à leur communication : les vocalisations impliquent des expressions oro-faciales, et les gestes pouvant être accompagnés de vocalisations (*e.g.* Arbib *et al.* 2008; Fröhlich *et al.* 2019; Gillespie-Lynch *et al.* 2014; Lemasson 2011; Levinson et Holler 2014; Liebal *et al.* 2014b; Masataka 2008; Meguerditchian et Vauclair 2014; Slocombe *et al.* 2011; Vigliocco *et al.* 2014). Ces différentes modalités de signalisation pouvant être perçues via des canaux sensoriels distincts (Liebal *et al.* 2014b), il est suggéré une possible complémentarité des fonctions permettant d'améliorer la transmission de l'information (Fröhlich et van Schaik 2018). Par exemple, les bonobos utilisent le même type de vocalisation (*i.e.* *contest-hoot*) dans deux contextes très distincts : le jeu et l'agression. L'ajout d'un geste communicatif permet de clarifier le message et d'éviter les ambiguïtés lors des interactions sociales (Genty *et al.* 2014). Des observations similaires chez les chimpanzés (Hobaiter *et al.* 2017; Wilke *et al.* 2017) mettent en évidence le fait que l'aspect multi-modal (*i.e.* geste associé à une vocalisation) précise la nature du signal transmis au receveur et lève de potentiels ambiguïtés sans pour autant être redondant (Fröhlich *et al.* 2019).

Cette théorie est aussi corroborée par des études neuro-anatomiques menées chez l'humain (*e.g.* Bernardis *et al.* 2008; Gentilucci et Volta 2008; Xu *et al.* 2009) qui suggèrent que les vocalisations et les gestes seraient sous le contrôle d'un même système neuronal localisé dans l'hémisphère cérébrale gauche. Un argument supplémentaire réside dans l'importance que semble avoir la communication gestuelle dans l'ontogenèse du langage oral (*e.g.* Vauclair et Cochet 2016).

Les pressions de sélection socio-écologiques ayant mené à l'émergence, au maintien et à l'élaboration des propriétés langagières chez les ancêtres communs aux primates actuels ont pu prendre différentes formes. Ainsi, il est souvent mentionné le cas du passage de l'arboricolisme à la terestrialité qui aboutit à une modification de la posture de locomotion et à une augmentation notable de la visibilité (*e.g.* Corballis 2009; Kendon 2004; McNeil 1992). De plus, la complexification de la vie sociale implique une plus grande diversité de rôles sociaux, différents niveaux de coopération et de compétition, de même que des innovations et des transmissions culturelles (*e.g.* Burkart *et al.* 2018; Dunbar 2003; Fitch 2010; Kaplan *et al.* 2009; Lou-Magnuson et Onnis 2018; Nowak et Highfield 2011; Pagel 2012) qui ont pu être des facteurs de sélection non-négligeables. Ainsi, en tenant compte d'un équilibre coût-bénéfice de la production d'un signal communicatif, l'impact de plusieurs facteurs influençant la communication des primates est à considérer à plusieurs échelles en fonction des caractéristiques propres (Prieur *et al.* 2020) : (*i*) à l'espèce (*e.g.* son écologie, système social,

morphologie/anatomie), (ii) à l'individu (*e.g.* les facteurs socio-démographiques, génétiques, physiologiques et psychologiques), (iii) aux contextes de production des signaux communicatifs (*e.g.* le type de contexte, la valence émotionnelle de ce dernier, l'effet d'audience), et (iv) aux signaux eux-mêmes (*e.g.* la modalité de production, les éléments qui composent le signal et leurs structures séquentielles, le contenu informatif et le niveau d'intentionnalité : Prieur *et al.* 2020). La prise en compte de l'ensemble de ces éléments amène à considérer une origine multi-causale du langage (Prieur *et al.* 2020).

Les probables origines évolutives du langage humain sont de plus en plus investiguées au travers d'approches couplant les différentes modalités de production des signaux communicatifs (*i.e.* approche multi-modale), tout en tenant compte des facteurs ultimes et proximaux influençant la production de ces signaux (*i.e.* approche multi-causale).

### 3. La place des études sur les interactions vocales dans le débat sur les origines du langage

Chez l'espèce humaine, la capacité à interagir vocalement apparaît très tôt au cours de l'ontogénie (*e.g.* Gratier *et al.* 2016) et se retrouve universellement chez toutes les populations (Stivers *et al.* 2009). L'apparition précoce de cette capacité langagière chez l'humain peut amener à se questionner quant à l'importance des interactions vocales comme base pour la complexification de la communication au cours de l'évolution (Takahashi *et al.* 2013). Toutefois, jusqu'à aujourd'hui, les hypothèses et théories sur les origines du langage se focalisent sur le comportement de l'émetteur ou du receveur, mais ignorent l'interaction dans son ensemble (Chow *et al.* 2015). Dans la partie qui va suivre, nous aborderons la grande diversité d'interactions vocales qui existe au sein du règne animal. Ces interactions se distinguent les unes des autres par leurs organisations et leurs fonctions.

Une interaction vocale est, par essence, un comportement social. Elle peut se définir sur deux dimensions distinctes mais complémentaires : une première purement temporelle et une seconde purement sociale.

En se basant sur la description des différents types d'interactions vocales rencontrées dans la littérature contemporaine (*e.g.* Yoshida et Okanoya 2005; Ravignani *et al.* 2014; Henry *et al.* 2015; Pika *et al.* 2018), il est possible d'en proposer une définition générale. Ainsi, une interaction vocale se manifeste quand au moins deux individus produisent, simultanément ou alternativement (*i.e.* vocalisations séparées par un court silence), une ou plusieurs vocalisations dans une fenêtre temporelle donnée. Néanmoins, cette définition temporelle ne suffit pas à

déterminer si ces productions vocales, provenant d'individus distincts, doivent être considérées comme une interaction à proprement parler ou non. En effet, plusieurs individus peuvent produire des sons indépendamment suite à la présentation d'un stimulus similaire (*e.g.* deux musiciens suivant un même chef d'orchestre), ou suite à l'influence de stimuli distincts survenant dans la même fenêtre temporelle (*e.g.* deux musiciens côte à côte, jouant simultanément deux partitions différentes). L'existence d'une interaction causale entre les émetteurs est essentielle pour que leurs productions vocales puissent être considérées comme constituant une interaction vocale (Ravignani *et al.* 2014).

Les interactions vocales peuvent prendre des formes diverses et sont ainsi classées selon différentes catégories en fonction de leurs organisations temporelles et/ou sociales. Trois principaux types d'interactions vocales se distinguent dans le règne animal : (i) les *chorus*, (ii) les duos et (iii) les échanges vocaux organisés sur le tour de parole.

### 3.1. Les *chorus*

Le terme de *chorus* fait référence à la production simultanée d'un même type de vocalisation par plus de deux individus (Yoshida et Okanoya 2005). Ce type de production vocale peut, à tort, être considéré comme une cacophonie sonore, à l'instar de la production de vocalisations d'alarme par un groupe suite à la détection d'un prédateur (*e.g.* cercopithèques à diadème : Papworth *et al.* 2008), bien que de nombreuses études révèlent l'existence d'une organisation temporelle sous-jacente (Hartbauer et Römer 2016; Ravignani *et al.* 2014). Des *chorus* ont pu être identifiés dans des clades phylogénétiquement distants (*i.e.* insecte, anouère, oiseau et mammifère : Burt et Vehrencamp 2005). Ces productions, impliquant des signaux sonores de longues portées (*e.g.* Christie *et al.* 2004; Faragó *et al.* 2014; Robinson et Hall 2002; Waser et Waser 1977), ont généralement lieu au crépuscule (*e.g.* Burt et Vehrencamp 2005; Robinson et Hall 2002). L'influence de facteurs abiotiques (*e.g.* température, hygrométrie et direction du vent) inhérents à la période de production et favorisant la propagation sonore peuvent être proposés (Brown et Handford 2003; Ey et Fischer 2009; Naguib et Wiley 2001; Waser et Waser 1977), mais il n'existe pas encore de consensus quant aux fonctions de cette démonstration acoustique, bien que plusieurs pistes non-exclusives soient proposées par les auteurs.

Les mâles de certaines espèces sont particulièrement enclins à produire des *chorus* lors de la période de reproduction. En fonction des stratégies, les chœurs favoriseraient la localisation du site de reproduction par les femelles des régions environnantes (*e.g.* anouères: Hylidae *Trachycephalus coriaceus* : Fouquet *et al.* 2020). Toutefois, cette superposition peut



masquer le signal acoustique individuel des prétendants, réduisant ainsi leur chance d’être perçus, et donc choisis comme partenaire par la femelle (*e.g.* Greenfield 2014). Dans ce contexte, les *chorus* peuvent être considérés comme le résultat de la compétition sexuelle inter-mâles (*e.g.* insectes : orthoptère *Neoconocephalus spiza* : Greenfield et Roizen 1993; Greenfield 1994), l’individu initiateur du *chorus* ayant plus de chance d’être perçu, et donc sélectionné par la femelle (*e.g.* insectes : Orthoptère : Greenfield et Roizen 1993; sauterelle *Mecopoda elongata* : Hartbauer *et al.* 2005; anoures : grenouille tungara *Engystomops pustulosus*, anciennement *Physalaemus pustulosus* : Greenfield et Rand 2000). En dehors de ces périodes de reproductions, des *chorus* résultant de la compétition territoriale entre les individus peuvent également être reportés (*e.g.* insectes : Orthoptère : Robinson et Hall 2002 ; anoures : crapaud accoucheur *Alytes obstetricans* et *Alytes cisternasii* : Bosch et Márquez 2000 ; Wells 1977 ; oiseaux : mésange à tête noire *Poecile atricapillus* : Foote *et al.* 2011 ; mésange charbonnière *Parus major* Kacelnik et Krebs 1982 ; mammifères : colobe *Colobus guereza* Schel et Zuberbühler 2012).

Chez certaines espèces vivant en groupes sociaux, les *chorus* peuvent se manifester entre les membres du groupe (*e.g.* oiseaux: cassican flûteur : Brown et Farabaugh 1991; mammifères : singe hurleur *Alouatta caraya* : Da Cunha et Byrne 2006; loups *Canis lupus* : Mazzini *et al.* 2013; chimpanzés : Mitani et Brandt 1994). Selon certains auteurs, ces vocalisations visent à prévenir les intrusions sur le territoire en maintenant les distances inter-groupes (*e.g.* mammifères : maki vari *Varecia variegata* et indri *Indri indri* : Geissmann et Mutschler 2006; loup : Harrington et David 1979), tout en défendant l’accès aux ressources sur une certaine zone géographique (*e.g.* mammifères : loup : Faragó *et al.* 2014; Harrington et David 1979). Les *chorus* peuvent ainsi être considérés comme un signal honnête de la « puissance défensive » du groupe résident à l’intention d’un groupe rival (*e.g.* mammifères : singe hurleur : Da Cunha et Byrne 2006), permettant d’éviter l’affrontement physique pouvant être très coûteux (Bradburry et Vehrencamp 1998; Kitchen *et al.* 2003). Il est aussi suggéré que ce type de comportement vocal faciliterait la cohésion et la coordination des membres du groupe dans un environnement visuellement obstrué (*e.g.* mammifères : loup : Faragó *et al.* 2014; maki vari et indri : Geissmann et Mutschler 2006).

En outre, certaines investigations révèlent que l’implication des participants à un même *chorus* n’est pas aléatoire, mais serait révélatrice des affinités sociales existantes entre les interlocuteurs (*e.g.* oiseaux : cassican flûteur : Brown et Farabaugh 1991 ; mammifères : chimpanzé : Fedurek *et al.* 2013a; loup : Mazzini *et al.* 2013). La participation à ces productions sonores favoriserait le maintien de liens sociaux forts (Fedurek *et al.* 2013a).

### 3.2. Les duos

Ce terme est traditionnellement utilisé pour qualifier les interactions vocales impliquant exclusivement deux partenaires de sexes opposés (Bailey 2003), exception faite des duos formés par les manakins fastueux mâles (*Chiroxiphia linearis*) lors des parades nuptiales (Trainer *et al.* 2002). Certains auteurs restreignent l'utilisation de ce terme à la synchronisation temporelle exacte de notes identiques (*e.g.* Langmore 2002), tandis que d'autres l'élargissent à la superposition des vocalisations des deux partenaires (*e.g.* Haimoff 1986; Hall 2004; Müller et Anzenberger 2002).

A l'image des *chorus*, les duos se retrouvent dans des taxons divers (*e.g.* insectes : Coleoptère, Hemiptère, Neuroptère, Orthoptère, Plecoptère : Bailey 2003 ; anoures : xénope du Cap *Xenopus laevis* : Tobias *et al.* 1998), et plus particulièrement chez les espèces dites monogames (*e.g.* oiseaux : cassican flûteur : Brown et Farabaugh 1991 ; Farabaugh 1982 ; ménis monias *Monias benschi* : Seddon *et al.* 2002 ; mahali à sourcils blancs *Plocepasser mahali* : Voigt *et al.* 2006 ; amazon à nuque d'or *Amazona auropalliata* : Wright et Dorin 2001 ; mammifères, surtout primates : *Hylobatidae* : Geissmann 2002 ; tarsier *Tarsius spectrum*, indri, langur de Mentawaï *Presbytis potenziani* et gibbons *Hylobates spp.* : Haimoff 1986 ; singe titi *Callicebus cupreus* : Müller et Anzenberger 2002). La grande majorité des vocalisations produites par les espèces lors des duos peuvent être considérées comme des vocalisations de longues distances au vu de leurs propriétés acoustiques (Seibt et Wickler 1977). Ainsi, chez les insectes et les anoures, l'une des principales fonctions des duos en période d'accouplement est la phonotaxie, induisant la localisation du congénère par l'émission de ces vocalisations et l'attraction (mutuelle ou non) des partenaires (*e.g.* insectes : *Barbitistes* : Stumpner et Meyer 2001; sauterelle *Poecilimon ornatus* : von Helversen *et al.* 2001; léptophye ponctuée *Leptophyes punctatissima* : Zimmermann *et al.* 1989 ; anoures : Tobias *et al.* 1998).

Chez les espèces monogames d'oiseaux et de mammifères, les duos auraient pour fonction de maintenir le contact entre les individus constituant déjà un couple (*e.g.* Wickler 1980) et/ou de synchroniser la physiologie des deux individus en vue de favoriser la reproduction (Schwabl et Sonnenschein 1992). La nature des vocalisations employées (*i.e.* vocalisations de longues distances) interpelle quant aux fonctions du duo qui, en plus de sa fonction au sein de la dyade, serait aussi adressé aux individus externes à celle-ci (Hall 2004; Seibt et Wickler 1977). Ainsi, le duo peut être interprété comme une forme de production vocale coopérative visant à défendre un territoire et/ou certaines ressources (*e.g.* oiseaux: gralline pie *Grallina cyanoleuca* : Hall 2000; usambiro barbet *Trachyphonus usambiro* et *T. d'arnaudii*

*emini* : Seibt et Wickler 1977), mais aussi à exprimer aux potentiels rivaux sexuels le statut dans lequel se trouvent les partenaires afin d'éviter l'usurpation de l'un des deux partenaires de la dyade (*e.g.* oiseaux: chouette hulotte *Strix aluco* : Appleby *et al.* 1999; Hall 2004). Dans ces différents contextes, en plus de renforcer la cohésion au sein de la dyade, la synchronisation des partenaires lors de l'émission du duo est cruciale et reflèterait le degré d'affinité existant entre les individus (Geissmann 2002; Hall 2004; Smith 1994; Wickler 1980).

### 3.3. Les échanges vocaux organisés sur le tour de parole

Ce type d'interactions vocales correspond à un échange ordonné de signaux acoustiques entre au moins deux individus, peu importe leur âge et/ou sexe (Pika *et al.* 2018; Yoshida et Okanoya 2005). Les échanges vocaux se caractérisent par un tour de parole, qui se traduit par une alternance de signaux suivant une certaine régularité temporelle et un évitement actif de leur recouvrement (Henry *et al.* 2015). Les échanges vocaux peuvent se manifester dans des contextes distincts et impliquer des vocalisations de faible intensité (*e.g.* Lemasson *et al.* 2010), ou des vocalisations de forte intensité, favorisées pour la communication sur de grandes distances (*e.g.* Schamberg *et al.* 2016).

Par exemple, en fonction de la stratégie de reproduction propre à l'espèce, l'émetteur va activement éviter de superposer son signal acoustique à celui de ses rivaux pour maximiser la perception du signal par le partenaire de sexe opposé. Il en résulte une alternance de signaux entre les mâles, permettant à la femelle de s'orienter par phonotaxie vers le mâle qu'elle a sélectionné (*e.g.* anoues : grenouille marbrée *Hyperolius marmoratus* : Grafe 1996; rainette crucifère *Pseudacris crucifer*, rainette versicolore *Dryophytes versicolor* et *Dendropsophus microcephalus* : Schwartz 1987).

L'intrusion d'un individu sur un territoire rival peut aussi être l'occasion d'observer la mise en place de ce type d'interaction vocale. L'individu résidant a tendance à répondre fréquemment à la vocalisation de l'intrus (*e.g.* anoues : Wells 1977 ; mammifères : daman du Cap *Procavia capensis* : Goll *et al.* 2017; chevreuil *Capreolus capreolus* : Reby *et al.* 1999), parfois en modifiant la structure acoustique de sa réponse (*e.g.* insectes : Orthoptère *Ligurotettix planum* : Greenfield et Minckley 1993 ; mammifères : chauve-souris *Saccopteryx biliata* : Behr *et al.* 2009). Le délai inter-vocalisations peut aussi être modulé, comme c'est le cas du troglodyte mignon (*Troglodytes troglodytes*) chez lequel le résidant réduit son délai de réponse à la suite de la vocalisation de l'intrus (Camacho-Schlenker *et al.* 2011). Il a pu être reporté que la superposition des signaux acoustiques peut mettre fin à l'interaction et aboutir à la fuite de l'individu « superposé » (*e.g.* oiseaux : étourneau : Henry *et al.* 2015; Naguib et Mennill 2010),

ou à l'augmentation du niveau d'excitation de ce dernier (*e.g.* rouge gorge *Erithacus rubecula* : Dabelsteen *et al.* 1997; mésange à tête noire : Mennill et Ratcliffe 2004). Chez la mésange à tête noire, ce sont les mâles dominants qui induisent le plus souvent cette superposition, reflétant un regain d'agressivité de leur part vis-à-vis de leur interlocuteur (Baker *et al.* 2012). A la vue de ces résultats chez les oiseaux, certains auteurs suggèrent que la superposition peut être perçue comme un signal en soit (Naguib et Mennill 2010), et peut-être plus précisément comme une agression dirigée vers l'interlocuteur (Naguib et Kipper 2006).

Toutefois, les échanges vocaux ne sont pas toujours le résultat d'une interaction agonistique entre les protagonistes. Bien que leurs fonctions ne soient pas précisément déterminées, il semble que ces interactions vocales puissent favoriser la localisation des congénères dans un environnement visuellement obstrué, faciliter la reconnaissance et les déplacements inter-groupes des individus isolés (*e.g.* mammifères : chauve souris *Diaemus youngi* : Carter *et al.* 2008; baleine franche australe *Eubalaena australis* : Clark et Clark 1980; grand dauphin : Quick et Janik 2012; bonobo : Schamberg *et al.* 2016). En outre, les vocalisations échangées entre les membres d'un même groupe peuvent aussi avoir pour objectif de réduire les potentiels conflits et de renforcer la cohésion sociale entre les interlocuteurs (*e.g.* mammifères : orque : Miller *et al.* 2004; cachalot *Physeter macrocephalus* : Schulz *et al.* 2008; éléphant *Loxodonta africana* : Soltis *et al.* 2005; rat-taupe nu *Heterocephalus galber* : Yosida *et al.* 2007; Yosida et Okanoya 2009). Néanmoins, des investigations supplémentaires sont essentielles pour déterminer la fonction sociale de ce type d'interactions (*e.g.* oiseaux : corbeau à gros bec *Corvus macrorhynchos* : Kondo *et al.* 2010; mammifère : beluga *Delphinapterus leucas* : Morisaka *et al.* 2013).

Les caractéristiques temporelles et sociales qui régissent leurs échanges vocaux, la flexibilité dont ils font preuve et les découvertes ontogénétiques récentes dans ce domaine ont conduit certains auteurs à lier l'étude de ce pattern d'interaction avec les théories sur l'évolution du langage et des conversations, et ce, chez différentes espèces (Henry *et al.* 2015). Par la suite, nous proposons une revue de la littérature contemporaine qui souligne les caractéristiques que partagent les conversations humaines et les échanges vocaux organisés des primates non-humains.



# Article 1

## Conversations among primate species

Loïc Pougault<sup>1,2,3</sup>, Florence Levréro<sup>2</sup>, Alban Lemasson<sup>1</sup>

<sup>1</sup> Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000 Rennes, France.

<sup>2</sup> Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES / CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France.

<sup>3</sup> ZooParc de Beauval & Beauval Nature, 41110 Saint Aignan, France

*Chapitre publié dans l'ouvrage **The Origins of Language Revisited : Differentiation from Music et the Emergence of Neurodiversity et Autism**, par N. Masataka. Editeur : Springer*

[https://doi.org/10.1007/978-981-15-4250-3\\_4](https://doi.org/10.1007/978-981-15-4250-3_4)

## Chapter 4

# Conversation Among Primate Species



Loïc Pournault, Florence Levréro, and Alban Lemasson

**Abstract** The literature in psychology and sociolinguistic suggests that human interlocutors, when conversing, virtually sign a sort of contract that defines the exchange rules in both structural and social domains. These rules make the messages more understandable and the interaction more predictable, but they may also act as a social bond regulator. These rules can be very basic such as speech overlap avoidance, respect of response delays, turn-taking and vocal accommodation to the context and interlocutor's social status. Interestingly, these rules are universally spread among human cultures questioning their biological basis and motivating the search for possible parallels with our primate cousins. Here, we will review the available literature on monkeys and apes. We will describe the different forms of vocal interactions, the temporal rules underlying these coordinated interactions, the non-random social selection of interlocutors and the context-dependent acoustic plasticity associated to these exchanges. The fact that primate species are socially varied, in terms of both social structure and social organisation, is another interesting aspect, since different social needs may predict different vocal interaction patterns and conversational rules. For example, duets, choruses and dyadic exchanges are not randomly distributed in the primate phylogeny and may even show different

---

Florence Levréro and Alban Lemasson contributed equally with all other contributors.

---

L. Pournault

Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) – UMR 6552, F-35000, Rennes, France

Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES/CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France

ZooParc de Beauval & Beauval Nature, Saint Aignan, France

F. Levréro

Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES/CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France

A. Lemasson (✉)

Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) – UMR 6552, F-35000, Rennes, France

e-mail: [alban.lemasson@univ-rennes1.fr](mailto:alban.lemasson@univ-rennes1.fr)

functions. Also, age proximity, kin membership, social affinity and hierarchy seem to play species-specific roles. Regarding plasticity, cases of vocal sharing and acoustic matching have been described in some species, notably in contact calls which are the calls the most frequently involved in dyadic exchanges. At last, a few studies also show that these ‘primitive’ conversational rules are often broken by juveniles and that the appropriate way to vocally interact with others may be socially learned, thus another aspect that do not seem strictly human.

**Keywords** Vocal accommodation · Conversation · Turn-taking · Vocal interaction

## 4.1 Introduction

The psychology and sociolinguistic literature suggests that human interlocutors, when conversing, virtually sign a sort of contract that defines exchange rules concerning both the structural and the social domains. These rules make messages easier to understand and interactions more predictable, and in addition they can act as a social bond regulator. These rules can be basic such as speech overlap avoidance, respect of response delays, turn-taking and vocal accommodation in relation to context and interlocutors’ social status. The fact that these rules are universal among human cultures questions their biological basis and motivates the search for parallels with our primate cousins. Here, we review the relative literature concerning monkeys and apes. We present different forms of vocal interactions, the temporal rules underlying coordinated vocal interactions, the non-random social selection of interlocutors and the context-dependent acoustic plasticity associated with these exchanges. The fact that primates live in various social systems in terms of both social structure and social organisation is another interesting aspect. Different social needs may predict different vocal interaction patterns and conversational rules. For example, duets, choruses and dyadic exchanges are not randomly distributed over the primate phylogeny and may even possess different functions. Also, age proximity, kin membership, social affinity and hierarchy seem to play species-specific roles. Plasticity in cases of vocal sharing and acoustic matching has been described for some species, notably for contact calls that are most frequently involved in dyadic exchanges. A few reports show that these ‘primitive’ conversational rules are often broken by juveniles and that the appropriate way to interact vocally with others can be learned socially, thus revealing an aspect that does not seem to be strictly human.



## 4.2 *Homo sapiens*' Conversational Rules

### 4.2.1 A Universal 'Contract of Communication'

Informal verbal interactions are the core matrix for human social life (Stivers et al. 2009). Anthropologists, ethnographers, psychologists, ethologists and sociolinguists agree that, even though notable cultural variations exist, comparable conversational basic phenomena appear in unrelated languages (Beach 1990). The universality of conversations has been well illustrated by the French social psychologist Rodolphe Ghiglione in his so-called 'contract of communication' (Ghiglione 1986). When interlocutors are engaged in a conversation, they virtually sign a sort of contract based on shared principles concerning both the temporal and social domains. This contract facilitates vocal coordination between interlocutors and promotes inter-comprehension. Ghiglione (1986) listed several principles, among which 'context relevance' (i.e. evaluation of context as pertinent or not to initiate a conversation), 'reciprocity' (i.e. evaluation of the partner as a valid interlocutor) and 'contract-based temporal rules' (i.e. respect of turn-taking between speakers and speech overlap avoidance).

#### 4.2.1.1 Temporally Ruled Interactions

Turn-taking represents the social core of human interactions (Levinson 2016). It is characterised by a reciprocal exchange of alternating, short and flexible turns between two or more interlocutors (Pika et al. 2018). Conversing is a well-known example of behaviour based on turn-taking (Sacks et al. 1974), and this key temporal coordination can be found in all languages and human societies around the world (Stivers et al. 2009). To engage in turn-taking, interlocutors A and B have to respect three basic rules: (i) A and B coordinate their speech temporally and avoid talking simultaneously; (ii) at any given moment during a conversation, interlocutors have coordinated roles – one of them sends a signal and the other responds (i.e. following the pattern AB and not AA or BB); (iii) during a longer conversation, the two interlocutors alternate (i.e. going from AB to ABA, ABAB, etc.) (Ghiglione 1986; Kerbat-Orecchioni 1990; Levinson 1983).

The optimal timing to take turns depends on the ability to coordinate language comprehension (prediction) and language production (formulation) (Sacks et al. 1974). Because, response delays are short, a speaker must begin formulating his/her response while his/her partner is still talking, and this act of formulation must be based on a prediction about what the partner is going to say (Sacks et al. 1974). The silent gaps between turns must last sufficiently to avoid overlap but be short enough to allow a certain fluidness during the verbal exchange. The brevity of silent gaps, often less than 200 ms, is particularly striking because linguistic formulation takes time: preparing to name a picture, for instance, takes at least 600 ms (Indefrey and

Levelt 2004) and can rise to about 1500 ms on average for a sentence (Griffin and Bock 2000).

Moreover, the timing of turn-taking is less precise and less homogeneously distributed among societies than often claimed (Heldner and Edlund 2010). Humans' common response delay varies with their culture (Stivers et al. 2009), and it can be influenced selectively and locally by social aspects of context (Sacks et al. 1974). For instance, the common response delays for subjectively on-time responses of Nordic cultures interlocutors (e.g. Danish 203 ms) are longer than those of Japanese (36 ms) (Stivers et al. 2009). These different response delays could be due to a specific cultural interactional pace or stem from more general differences of the overall tempo of social life (Hall 1959). Despite these differences, the available data are consistent with the presence of a universal and stable system of turn-taking avoiding overlap and minimising silent gaps (Stivers et al. 2009).

#### 4.2.1.2 Socially Guided Interaction Patterns

Human oral communication is typically intentional and directional, targeting a dedicated audience or a particular individual interlocutor. To engage in a conversation, interlocutors are thus not chosen randomly. Affinity and age proximity with a chosen interlocutor are often considered determining factors in both traditional and modern human societies (Bascom 1942). However, humans do not converse exclusively with friends. Humans are hence able to adjust their social distance to the interlocutor during a given conversation due to an acoustic convergence/divergence phenomenon named 'vocal accommodation' (Babel 2010; Chartrand and van Baaren 2009; Gallois et al. 2005). Short-term accommodation (i.e. minutes or hours) can be distinguished from long-term accommodation (i.e. month or years), and long-term accommodation is commonly assumed to be based on repeated short-term accommodation (Nguyen and Delvaux 2015; Trudgill 2008). The communication accommodation theory claims that linguistic short-term convergence and its opposite, divergence, are strategically used by speakers engaged in spoken interactions in order to respectively minimise or maximise their social distance with their interlocutors, so as to reinforce their own social identity (Gallois et al. 2005). In this way, speakers are able to match various kinds of acoustic features, such as fundamental frequencies or vowel spectra, in response to social factors such as attraction (Babel 2012) and strength of the relationship (Babel 2010; Pardo et al. 2012). During long-term accommodation, the change observed in speech-involved attributes is measured over long stretches of dialogue such as accent, speaking rate, intensity, low-frequency band variation, pause frequency and utterance length (Gregory 1990). Convergent characteristics in vocal behaviour reflect a speaker's motivation to increase social integration or identification (McGarva and Warner 2003).

### ***4.2.2 Functions of Conversations***

Humans may converse to transmit a message to a specific person. They can develop different strategies in order to increase the chance of success to continue a given conversation, such as persistence (i.e. signal repetition) and elaboration (i.e. signal variation) (Golinkoff 1986, 1993). However, observational studies of human conversations in relaxed social settings suggest that these consist predominantly of exchanges of social information (mostly concerning personal relationships and experience) (Dunbar et al. 1997). Conversations are thus used to facilitate social integration and social bonding.

Respect of turn-taking during conversation, ‘as an orderly distribution of opportunities to participate in social interaction’, is considered one of the ‘most fundamental preconditions’ for a viable social organisation (Schegloff 2002). Apart from just ‘politeness’, maintaining mutual comprehensibility when participants talk at the same time is obviously difficult (Duncan 1972). When this occurs, overlap is mostly associated with negative and strong personality attributes (Ter Maat et al. 2010) and considered as serious impoliteness by most human societies (Calame-Griaule 1965). There is a fundamental distinction between accidental overlap and deliberate overlap (Kurtić et al. 2013). Conversely, pauses between turns are considered as more agreeable (Ter Maat et al. 2010). Nevertheless, long silence gaps are considered less affiliative, or more distancing, than shorter ones (Roberts et al. 2011). Temporal rules make messages more understandable and interactions more predictable, but they can also act as a social bond regulator.

### ***4.2.3 Ontogeny of Conversational Behaviours***

Maturation and social experience clearly shape efficiency in turn-taking, but turn-taking is definitely found at very early ages. Reissland and Stephenson (1999) compared vocal interactions between mothers and their babies who were either premature (range 26–32 weeks postmenstrual, mean = 30.2 weeks) or term infants (from 38 to 42 weeks postmenstrual, mean = 40.1 weeks). Their conversational turns were investigated at the hospital (1 week old for term infants and 39 weeks postmenstrual for preterm infants) and then at home 2 months later. No differences in vocalisation rates or conversational turns could be evidenced at the hospital. However, when they were 2 months old, mothers of preterm infants responded more to their infants’ non-cry vocalisations, and they took every opportunity to do so. Term infants were more likely to respond to their mother’s vocalisations. Another report focusing on mother-preterm conversational turns found that both adult word count per hour and infant vocalisation count per hour increased from 32 to 36 weeks postmenstrual age (Caskey et al. 2014). When a parent was present, infants produced significantly more conversational turns per hour at both ages.

Around 3 or 4 months old, infants are more capable to control their vocal production voluntarily. Two structurally different patterns of dyadic vocal interactions between mother and infant emerge: the ‘coaction mode’ when both members perform the same or similar behaviour at the same time and the ‘alternating mode’ when mothers often create or shape an antiphonal or alternating pattern of vocalisations between themselves and their infants (Stern et al. 1975). While more overlaps than alternations are found during the second semester of life, the contrary is observed during the third semester (Ginsburg and Kilbourne 1988).

Caregivers play a major role in the development of conversational abilities. They use cues such as stable timing and exaggerated pauses to help infants learn to take turns (Stern et al. 1977; Stern and Gibbon 1979). Longitudinal approaches reveal that the structure of vocal interactions elicited by the mother’s behaviour is temporally stable during the first half year of life in face-to-face interactions (Cossette et al. 1986; Stern et al. 1977). At very early stages –before 2 months – infants’ vocal production occurs in the form of periodically occurring bursts (Schaffer 1977). Between these bursts, the mother skilfully inserts her own actions, and by doing so, she is attempting to provide some temporal organisation for her interaction with the infant. Mothers’ speech to 2-month-old infants is timed in such a way as to leave time for the infant’s response. Maternal utterances are brief (between 0.5 s and 1.5 s) followed by pauses lasting around 1 s and usually followed by another utterance. Pauses between utterances that are connected through repetition of form, content or topic rarely exceed 3 s and pauses longer than 3 s generally demarcate the end of mutual engagement episodes (Stern et al. 1977; Stern and Gibbon 1979). The majority of maternal responses to infants aged between 2 and 4 months occur within 1 s after the signal (Keller et al. 1999) and pauses between turns of young infants and mothers range from 500 ms to 1 s (Jaffe et al. 2001).

Of course both caregivers and infants are actors in the acquisition process. The fact that mothers and infants engaged in a social interaction match each other’s duration of vocal utterances and pauses suggests mutual regulation of turn-taking (Beebe et al. 1985; Jaffe et al. 2001). This is in line with the theory of social learning that presupposes that infant vocalisations are reinforced by continuous social stimulation. Masataka’s (2003) experimental results support the anecdotal evidence that infants between 12 and 18 weeks old are less likely to begin vocalising while their caregivers are speaking than when they are silent (Ginsburg and Kilbourne 1988). Even when mothers stimulate their infants randomly, they maintained their infants’ attention, and infants focused their attention on their mothers to respond contingently. Neglected children fail to develop this ability to converse, showing irrelevant turns, interruptions, simultaneous talking and non-contingent responding. Moreover, vocal matching ability appears early during ontogeny. It occurs in 41–57% of infants’ non-crying vocalisations, during spontaneous vocal interactions between mothers and their 2-, 3- and 5-month-old infants. With age, matches become more and more complex in number and types of included features such as pitch, pitch modulation, duration and rhythm (Papoušek and Papoušek 1989).

These studies show that prelinguistic infants engage in proto-conversational interactions with their caregivers and that these interactions are well coordinated

and well timed (Hilbrink et al. 2015). However, when children begin to interact using language, their fluency at turn-taking declines; gaps in linguistic turn-taking are then longer than previously in nonlinguistic turn-taking (Hilbrink et al. 2015). This development trajectory, from well-timed nonlinguistic interactions to less fluent linguistic interactions, raises the possibility that preschool children (around 9 months old) may in fact lack competence in flexibly coordinating linguistic prediction and formulation (Lindsay et al. 2019). This slowing down occurs when the infants begin to grasp the significance of intentional communication. Levinson (2016) argued that the poor timing of children's conversational turn-taking is not indicative of an underlying lack of competence, but he suggests that children can flexibly coordinate prediction about what a conversational partner will say by formulating early a response, but compared to adults, it is hard for them to encode linguistically the ideas that they want to express, and thus they respond after a longer delay.

#### ***4.2.4 Underlying Neurobiological Processes***

The cognitive processes governing human behaviour during a verbal interaction are complex and vary with context. We chose to focus on one well-studied conversation context, the face-to-face interaction. Unlike other types of communication, face-to-face interactions involve more continuous turn-taking (Wilson and Wilson 2005). Recently Stephens et al. (2010) found that, when the speaker's voice is aurally presented to the listener, the brain activities of both listener and speaker synchronise. A successful communication notably results in a temporally coupled neural response pattern, a pattern not systematically found if speakers talk a language unknown to listeners. Using the near-infrared spectroscopy developed by Cui et al. (2012) that measures simultaneously the brain activity of two conversing persons, Jiang et al. (2012) examined neural mechanistic features of face-to-face communication compared with back-to-back dialogue, face-to-face monologue and back-to-back monologue. These authors highlighted a simultaneous neural synchronisation in the left inferior frontal cortex – left IFC – (usually important for language production) increased in face-to-face communication. This facilitation could be explained by an action-perception system (Fogassi and Ferrari 2007; Nishitani et al. 2005). The left IFC, in addition to other brain regions, is indeed a site where mirror neurons are located (Rizzolatti and Arbib 1998). These neurons are well known to respond to the observations of an action, to a sound associated with that action or to observations of mouth-communication gestures (Ferrari et al. 2003; Kohler et al. 2002).

### 4.3 Nonhuman Primates' Patterns of Vocal Interactions and Conversational-Like Features

The first part of this chapter shows that, despite obvious cultural variations, language seems 'by nature' interactive. This questions the existence of possible underlying biological bases and encourages scientists to tackle the evolutionary origin of 'conversational' behaviour. For some authors, the ability to converse in a rule-governed way with a particular interlocutor probably appeared somewhere before the common ancestor of modern humans and Neanderthals (600,000 years ago), since all the genetic and physiological prerequisites for speech seems in place (Dediu and Levinson 2013), while they seem lacking in *Homo erectus* (ca 1.6 My) (Levinson and Holler 2014; MacLarnon and Hewitt 2004). During the intervening million years, simple alternation of vocal utterances during vocal exchanges may have provided the framework for linguistic complexity to evolve (Levinson 2016). The temporal properties of turn-taking may have remained fixed, creating the foundations on which more complex linguistic material was progressively added. For instance, more complex linguistic material could be free flowing, intricately varied, rapid sound sequences suitable for the fast transfer of complex, highly flexible communication allowed by fine neurobiological control of the respiratory muscles and complex cognitive processes. Now, language diversity is considered as driven by cultural evolution (Levinson 2016). However, other authors suggest that the rules governing humans' vocal interactions are much more ancestral as reports stress parallels with birds and mammals (Henry et al. 2015) and particularly with nonhuman primates (Lemasson et al. 2011a, b; Snowdon and Cleveland 1984; Sugiura and Masataka 1995).

#### 4.3.1 Definition of Conservation-Like Vocal Exchanges

The vast majority of nonhuman primates are highly social species that live in groups governed by complex relationships (Dunbar 2009; Grubb et al. 2003; Kappeler and Van Schaik 2002; McComb and Semple 2005). Their social life requires the existence of communication rules respected by all individuals, allowing each individual to share information and to develop and to maintain social bonds with particular group members. Although the level of ability to communicate visually (mimics, gestures, postures) varies greatly among species, all nonhuman primates are vocal communicants (Lemasson 2011). In this context, vocal interactions play a key role in their social communication network (Pika et al. 2018).

Nonhuman primates produce a variety of call types, some dedicated to specific contexts, such as alarm calls, distress calls, copulation calls and food calls. These calls increase the chances of individuals to survive and to reproduce. However, several species also produce so-called 'social/contact' calls, not specifically linked to a given context but typically associated with peaceful and relaxed behaviours. They

are produced with high call rates (most often higher than the other call types) and usually involved in socially and temporally coordinated vocal exchanges. We believe that these contact call exchanges are good candidates for ‘conversation-like vocal interactions’ in primates. They involve a diversity of recurrent vocal partners. Interlocutors are typically familiar individuals belonging to a given social group and can be of any age and sex. Production of conversation-like vocal exchanges is not related to ‘primary need’ contexts, time of day or year, and they relate more to daily-basis short-distance communication. Thus, these types of interactions differ from vocal interactions related to reproduction in insects (Bailey 2003), amphibians (Forester and Harrison 2008) and even mammals like rodents (Okobi et al. 2019) or deer (Clutton-Brock and Albon 1979; Reby and McComb 2003).

To what degree vocal nonhuman primates’ communication is intentional is a matter of current debate. Nevertheless, the debate concerning vocal production (control of acoustic structure of the sound emitted) is more important than that concerning vocal usage (control of utterance of a vocalisation). Several reports show that vocal usage is socially aware. Indeed, the identity of potential receivers (i.e. active interlocutors or passive listeners) influences the vocal behaviour of emitters who can choose to vocalise only when particular conspecifics are present. Audience effect on call rates (but even on call structure to a lesser extent) has been found in a range of primate species. Good evidence shows that the composition of the nearby audience can affect the likelihood of calling. For example, some contact calls are preferentially emitted when approaching specific partners: baboons’ grunts and macaques’ girneys addressed to affiliated individuals (Silk 2002) and female/low-ranking male chimpanzees’ pant-grunts addressed to the higher-ranking male (Laporte and Zuberbühler 2010). However, the audience effect has been observed in a broad range of contexts and for several types of calls (e.g. chimpanzees: Leavens et al. 2004, 2009; Slocombe et al. 2010; macaques: Hauser and Marler 1993; brown capuchins: Pollick et al. 2005; red-capped mangabeys: Bouchet et al. 2010). The social and temporal rules underlying contact call exchanges are good examples to illustrate the level of control of vocal usage by nonhuman primates.

Conversation-like call exchanges are only one type of nonhuman primates’ vocal interactions. Authors typically distinguish this behaviour from other vocal patterns like isolated calling (one call is emitted and no other calls can be heard around), repeated calling (the same callers call several times in a row), disorganised phonoresponses (one individual produces a call, typically an alarm call, and all the group members respond in an apparent chaotic way), chorusing (two or more individuals overlap their emissions of a given call type) and duetting (two individuals synchronise long series of calls or songs with a stereotyped temporal association). Yoshida and Okanoya (2005) hypothesised that during evolution choruses came first, then duets and vocal exchanges finally emerged. These different calling patterns are however found at different levels of the primate phylogeny, and all the great apes present all these patterns. The two ‘non-conversational’ interaction patterns (i.e. chorus, duet) suggest abilities of animals to control the timing of their vocal utterances with predictable temporal coordination. For example, male chimpanzees



produce a species-typical long-distance call known as pant-hoot. This vocalisation is often emitted during choruses in a food arrival context (Clark and Wrangham 1993) or to maintain contact over long distances (Mitani and Nishida 1993). Pant-hoots start with a build-up introduction that facilitates synchronisation of the climax part of the pant-hoots of the different overlapping callers (Fedurek et al. 2013). Duetting generally concerns vocal interactions between two partners of the opposite sex (Bailey 2003) and are common among nonhuman primates with stable monogamy and territoriality (e.g. tarsiers, indris, Mentawai langurs and gibbons: Fuentes 1998; Geissmann and Orgeldinger 2000; Haimoff 1986; MacKinnon and MacKinnon 1980). Each species of gibbon has its species-specific temporal duet organisation with predictable male/female/couple parts (Geissmann 2002).

### ***4.3.2 Temporal Organisation of Conversation-Like Vocal Exchanges***

Campbell's monkeys' conversation-like vocal exchanges have been well studied. The adult males and females of this species form harem groups and live in West African forests. They present different vocal repertoires and patterns of vocal usage. Males typically produce long loud calls to warn about danger that trigger disorganised phono-responses from all females using their own versions of alarm calls (Ouattara et al. 2009a, b). However the social core of intra-harem communication relies on female contact calls. Their vocal interactions are organised in temporally structured ways generally involving only two or three interlocutors responding to each other (Lemasson et al. 2010) and forming the so-called contact call exchanges respecting a turn-taking pattern (Lemasson et al. 2011a). Exchanging partners can be physically close (e.g. grooming) or distant (without any visual contact). The respect of the turn-taking during vocal exchanges has been observed in a broad range of nonhuman primate species (e.g. pygmy marmosets: Snowdon and Cleveland 1984; common marmosets: Takahashi et al. 2013; Japanese macaques: Sugiura 1998; gorillas: Lemasson et al. 2018; bonobos: Levréro et al. 2019). A given caller typically produces a single call and waits for the vocal response of another group member before possibly calling again.

The first important rule observed is that the respondent adjusts the timing of its response by leaving a short silent gap after the initiator's call. This gap respects a maximum duration (about 1–3 s according to species) to ensure call coordination and a minimum silent delay to prevent call overlap (Henry et al. 2015). A study focusing on long periods of common marmosets' dyadic vocal interactions showed vocal cooperation and temporal coordination (Takahashi et al. 2013). Marmosets adjusted their vocal output reciprocally and continuously with each other. Dyads engage in turn-taking by showing that they prevent overlapping call (waiting ~5 s before responding); the timing of their calls is periodically coupled, and they mime each other so that when one speeds up or slows down its call timing, the other will follow.



Individual Campbell's monkeys respond to each other with a short intercall interval of up to 1 s but rarely less than 260 ms (Lemasson et al. 2010). The common response delay is rarely shorter than an average call duration (e.g. Diana's monkeys: Candiotti et al. 2012; squirrel monkeys: Masataka and Biben 1987; Japanese macaques: Sugiura 1993; spider monkeys: Briseno-Jaramillo et al. 2018; gorillas: Lemasson et al. 2018; bonobos: Levréro et al. 2019).

Response delay is however flexible as it varies with context, such as the physical distance between callers. The interval between Japanese macaque interlocutors is longer when group members are scattered, and more importantly, an initiator persists and repeats its call when it fails to get a response, and the farther away the targeted respondent, the longer it waits (Sugiura 2007). Sugiura (1993) also suggested a possible cultural effect on the common response delay as it varies among Japanese macaque' population. Thus, delays are longer in the Yakushima population than in the Ohirayama population. As the original members of the Ohirayama group were translocated from Yakushima island 33 years before this study, this difference suggests possible trans-generational social transmission of the conversational rule in addition to a dialect-like structural modification of the contact (coo) calls recorded in these two populations (Tanaka et al. 2006). Social determinants are also key factors that influence duration of the silence gap between squirrel monkeys' exchanged vocalisations. This gap is significantly shorter when the respondent has a strong affiliative bond with the initiator (Biben et al. 1986; Masataka and Biben 1987). Some species, such as common marmosets (Yamaguchi et al. 2009) and spider monkeys (Briseno-Jaramillo et al. 2018) can exchange two types of contact calls, i.e. marmosets' phees and trills or spider monkeys' whinnies and high whinnies. Adult interlocutors typically match the call type used to respond to one another. Nevertheless, response delay can be call type – dependent. Common marmosets respond faster to a trill vocalisation than to a phee vocalisation, and the delay between repeated trills, a way to elicit a response from conspecifics, is shorter (3 s) than the delay between repeated phees (6 s). Authors suggested that phees are usually longer and louder than trills and would thus be used preferentially by distant interlocutors (Yamaguchi et al. 2009)

Several studies furthered their investigations and showed that the respect of the expected temporal pattern matters for animals, at least adults. Playback experiments showed that adults discriminate spontaneously between artificial vocal exchanges respecting the turn-taking principle (Campbell's monkeys: Lemasson et al. 2011a) and the non-overlap principle (gorillas: Pougnault et al. 2020) and those breaking these rules. In both cases, subjects lost interest in the broadcast when the interaction did not respect the temporal rule. Also, in Japanese macaque, adults that produce repeated calls in a row, leaving no chance for turn-taking responses, are low-ranking males (Lemasson et al. 2013a, b).

These temporal features (i.e. turn-taking and overlap avoidance during vocal interactions) by themselves are not sufficient to distinguish primates' vocal interactions from those of other animal species as similar temporal features have been observed for a broad range of nonprimate species (insects: Hartbauer et al. 2005, Hedwig 2006; amphibians: Schwartz 1987, Zelick and Narins 1985; birds: Hyman

2003, Mennill et al. 2003, Vehrencamp et al. 2014; mammals: Behr et al. 2009, Carter et al. 2008, Ghazanfar et al. 2002, Goll et al. 2017, Miller and Wang 2006). What makes the vocal exchanges of some primate species particular is that they present the necessary combination of temporal and social dimensions that define ‘conversation-like’ vocal exchanges.

### ***4.3.3 Social Organisation of Conversation-Like Vocal Exchanges***

The ‘social bonding hypothesis’ proposed by Dunbar (1996) highlighted the key role played by oral interactions during the evolution of primates. This theory suggested that, in large social groups and in groups with solid bonds, the function of gestural grooming was partially transferred to vocal interactions. Gestural grooming plays an important role in forming new bonds and strengthening existing ones in all primate species (Dunbar 1991). As group size and bonding network increased, conversation-like oral exchanges may have developed since it then became impossible to allocate sufficient time to groom all partners physically, so vocal exchanges could play a ‘grooming-at-distance’ function (Dunbar 1996).

In line with Dunbar’s predictions, two recent studies confirmed that preferred contact call exchange partners are the same as physical grooming partners and that conversation-like exchanges can be considered grooming-at-distance (Japanese macaques: Arlet et al. 2015; spider monkeys: Briseño-Jaramillo et al. 2018). Moreover, female Japanese macaques that groomed each other more also responded to each other’s calls more. Thus, the exchange partner, as for humans, is not randomly nor opportunistically chosen. These findings are supported by several other studies reporting that individuals involved in a positive relationship are more likely to exchange calls (e.g. cotton-top tamarins: Jordan, et al. 2004, pygmy marmosets: Snowden and Cleveland 1984, bonobos: Levréro et al. 2019, Campbell’s monkeys: Lemasson et al. 2006, squirrel monkeys: Soltis et al. 2002).

Interlocutors can be selected according to other factors, such as social status. For Campbell’s monkeys, age of individuals is an important regulating factor; elders, regardless of dominance status, elicit more responses from their younger conspecifics despite a lower call production (Lemasson et al. 2010). The hierarchy of dominance in this species is very subtle, and agonism is rare (Lemasson et al. 2006). The attention paid to the age of the interlocutors has also been found for marmosets (Chen et al. 2009). Interestingly an effect of age has been found concerning western lowland gorillas’ conversation-like interactions. The closer two gorillas are in age, the more likely they are to exchanges soft calls (Lemasson et al. 2018). Age proximity could play a key role in great apes that are characterised by a long development and longevity (Levréro et al. 2019). Dominance status can matter more for other species. For example, dominant male and female white-faced

capuchins receive more vocal responses than subordinates of either sexes (Digweed et al. 2007).

#### ***4.3.4 Neurobiological Processes Associated with Vocal Exchanges***

Although much is known about the acoustic structure of primate calls and the social context in which their vocalisations are usually uttered, our knowledge concerning the neocortical control of vocal interactions is still limited. Information concerning this domain comes mainly from studies of lesions of squirrel monkeys, marmosets or macaques (Aitken 1981; Jürgens et al. 1967; Jürgens and Ploog 1970; Sutton et al. 1974). Nonhuman primates have been thought, for several decades, to have no voluntarily control over their vocal production, a fundamental difference from human language (Jürgens 1995). The first neurological studies on this topic showed that the production of some of squirrel monkeys' vocalisations was related to the limbic system and processed in subcortical areas associated with emotions. Nobody seriously disputes the fact that humans have far greater control over their vocal production apparatus and the evidence that some Old and New World monkeys and apes have limited vocal control (Lemasson et al. 2013a, b).

One line of evidence comes from studies based on conversation-like vocal exchanges. For example, when a marmoset hears a conspecific vocalisation and decides not to answer, the primary auditory cortex is activated (Simões et al. 2010). In contrast, when the individual answers, the ACC (i.e. anterior cingulate cortex), the VLPFC (left ventrolateral prefrontal cortex; note that the VLPFC is a key region for speech control in humans) and the DMPFC (dorsomedial prefrontal cortex) are all activated, suggesting the existence of context-dependent vocal control in nonhuman primates (Jürgens et al. 1996; Simões et al. 2010). Other studies reported activation of the ventral premotor cortex when a subject produces conditioned vocalisations, but not when a subject responds spontaneously to the presence of food (pigtailed macaques: Coudé et al. 2011). Single-cell recordings in the ventrolateral prefrontal cortex of monkeys trained to vocalise (contact calls) in response to visual cues evidenced call-related neurons (rhesus macaques: Hage and Nieder 2013).

A recent study showed a robust correlation between marmosets' cortical activity and conversation-like contact call exchanges and in addition that the magnitude of neuronal responses increased in relation to conversation length. This supports the notion that the neural process is strongly related to the social context of the vocal exchange (Nummela et al. 2017). Authors hypothesised that the change in the frontal motor cortex may reflect a change in social arousal and attention (Miller et al. 2010). This could serve as a sensory gating function to facilitate rapid processing of conspecific vocalisations throughout the auditory system (Miller et al. 2010) and precipitate a cascade of subsequent social decision-making processes (Toarmino et al. 2017).

Additionally, recently Okobi et al. (2019) identified a region of singing mice's motor cortex (the orofacial motor cortex) that influences vocalisation and mediates rapid vocal interaction. Authors showed that the motor cortex is required for adaptive counter-calling but not for the production of the call itself. This is the first demonstration of cortical dependence of a precise vocal interaction in mammals. Also, they demonstrated that the motor cortex is able to adjust dynamically the pacing and duration of call sequences, consistent with changes of these parameters during social interactions.

All these results support the notion that diverse cortical structures are involved in the control of nonhuman primates' vocal communication. Neuronal processes, such as the social context change in the frontal cortex state, may occur only when primates are actively interacting with one another.

### ***4.3.5 Acoustic Plasticity Associated with Vocal Exchanges***

As stated above, most human conversations deal with exchanges about personal relationships and experience (Dunbar et al. 1997). Interestingly, contact calls are often the call types that present the highest level of acoustic identity coding (Lemasson and Hausberger 2011 but see Keenan et al., sub). Beyond identity, these call types can also code for different sorts of social information through a process of vocal convergence. Over more or less long periods of time, individuals can modify the frequency modulation pattern of their contact call to copy their preferred partners (e.g. pygmy marmosets: Snowdon and Elowson 1999). Female Campbell's monkeys that are more closely affiliated are more likely to share particular vocal variants than conspecifics that are not, regardless of genetic relatedness (Lemasson et al. 2011b). However, an individual's variants are not stable over its entire adult life. Reports show that these changes are triggered by alterations of females' social relationships (Lemasson et al. 2003; Lemasson and Hausberger 2004). Authors showed, using playback experiments, that only current variants (conversely to older ones) elicited immediate call responses (Lemasson et al. 2005). Bonobos, another so-called tolerant species, similarly share variants (Levréro et al. 2019). The greater the age difference between two bonobo callers, the less likely they will share acoustic variants. These authors hypothesised that juvenile bonobos display a 'youngster vocal badge' as they often play with peers and this could favour the development of social bonds during their long immaturity period. However, vocal copying by Japanese macaques, who also often exchange contact (coo) calls but form despotic societies, does not explain social affinity. Instead, subordinates copy dominants' calls (Lemasson et al. 2016). Interestingly, macaque species vary socially, some are tolerant, while others are despotic (Thierry et al. 2000; Thierry 2007). Recently a study confirmed that hierarchy explains acoustic similarity in another despotic species, rhesus macaques, but not in their tolerant cousins, Tonkean macaques (De Marco et al. 2019).

At the time scale of a vocal exchange, acoustic matching can be observed in nonhuman primates. The respondent produces a call that matches the acoustic characteristics of the call used by the initiator, a phenomenon comparable to the vocal accommodation mentioned previously for humans. This has been clearly demonstrated using observations and playback experiments with Japanese macaques (Sugiura 1993, 1998) and bonobos (Levrero, pers. comm.). Female Diana monkeys converge vocally in social cohesion contexts and diverge in spatial cohesion contexts (Candiotti et al. 2012).

These copied acoustic variants can be seen as social badges that individuals use during vocal exchanges to advertise their affinity and to facilitate social integration (Lemasson and Hausberger 2004; Sewall et al. 2016).

Apart from convergence/divergence processes, acoustic plasticity can be found in conversation-like contact call exchanges. During vocal interactions, individuals can use several kinds of signals to elicit vocalisation by their interlocutors or to indicate their desire to interact vocally. Biben et al. (1986) demonstrated that during squirrel monkeys' vocal interactions, low-frequency chucks are more likely to initiate vocal exchanges, whereas high-frequency chucks terminate them. Call combination can be used for similar purposes. Red-capped mangabeys typically add an optional 'Uh' unit to their contact call when they engage in vocal exchanges (Bouchet et al. 2010). Bonobos also use call combination of long-distance vocalisations during interparty movements (Schamberg et al. 2016). When wild bonobos are motivated to move from one foraging subgroup to another, they produce a combination of 'whistle + high-hoot'. They are likely to travel to the other subgroup if their initial vocalisation elicited an answer; otherwise they demonstrate an apparent persistence in the production of 'whistle + high-hoot'. Moreover, individuals can modify the acoustic structure of their call combinations to allow the listener to distinguish between those given spontaneously and those given in a response to another call.

In addition, when a Japanese macaque fails to elicit a vocal response from its conspecifics, he/she can elaborate by emitting a second (acoustically modified) call to emphasise his/her motivation to receive a response (Koda 2004). The second call will be longer with frequency modulations of larger amplitude. Playback experiments confirmed that such modified calls trigger systematically more responses. Squirrel monkeys' (Biben et al. 1986) and common marmosets' (Choi et al. 2015) calls with 'exaggerated acoustic patterns' are used preferentially when the initiator is far from the targeted respondents.

### ***4.3.6 Breaking the Conversational Rules by Nonhuman Primates***

As in humans, the appropriate way to behave vocally when conversing results from social development in several nonhuman primate species. Immature individuals break social and temporal rules more often than adults, supporting the idea of a

possible learning process. For example, immature Japanese macaques and Campbell's monkeys are less likely to adhere to the turn-taking rules than adults (Lemasson et al. 2011b, 2013a, b). Analyses of spontaneous vocal exchanges of both species show that juveniles break the turn-taking rule more often (up to 12 times more frequently) by overlapping conspecifics' calls or by producing several repeated calls. Using playback experiments, these authors showed that, conversely to adults, juvenile Campbell monkeys do not detect or do not care about respecting (or not) turn-taking in these artificial call exchanges (Lemasson et al. 2011a).

Chow et al. (2015) suggested that turn-taking during common marmosets' vocal interactions is learned during ontogeny and that its development is guided by parents' behaviours. These authors found that parents were significantly less likely to produce a response following an interruption of their own call by their offspring than during vocal interactions when no interruption occurred. Interestingly, overlaps during vocal interactions are more frequent between offspring and their parents than between siblings. Additionally, despite only small differences in the duration of phee calls produced by mothers and fathers, offspring are significantly more likely to interrupt their father than their mother. Indeed, whereas juvenile marmosets adopt an adult-like temporal pattern of vocal exchanges with their mothers around 10–12 months, this is not the case during vocal interactions with their fathers. However, while vocal exchanges with mothers and fathers gradually slowed over the first year of life, the temporal pattern stayed relatively constant during sibling vocal interactions. Chow et al. (2015) suggested that these ontogenic changes are not simply due to a general maturational process but that marmosets learn to adapt the relative timing of their vocal responses based on the specific social context and the identity of their interlocutors.

The hypothesis presented by Chow et al. (2015) could be valid if the quantity of contingent parental responses was correlated with the rate of turn-taking development so that infants with more responsive parents learned to adjust the temporal pattern of their turn-taking faster. Takahashi et al. (2016) investigating parental influence on juvenile marmosets' turn-taking development could not evidence any relationships between maturation rate of vocal turn-taking and overall frequency of contingent parental responses. However, the capacity to grasp the vocal turn-taking temporal pattern increases with age of individuals. During the first post-natal week, marmosets interact only weakly with their parents, but these interactions increase with experience. However, the vocal pattern used by parents is not modified during this period. Takahashi et al. (2016) hypothesised that this was due to self-monitoring by young marmosets rather than parental influence.

A similar developmental pattern is observed for at respect of social and acoustic rules. For example, a recent study showed that free-ranging spider monkey adults respond preferentially to friends and perform systematically call-type matching immature did not and young individuals even less than older immatures (Briseno-Jaramillo et al. 2018). Adults may have either the higher cognitive abilities required for processing such complex information or the social experience necessary for understanding these conversational rules. Subadult Japanese macaques (aged 3–5 years old) were less efficient than adults at respecting the call-matching rule during

interindividual exchanges of coo calls (Masataka 2003, Sugiura 1998, Sugiura and Masataka 1995). A playback study ran by Bouchet et al. (2017) confirmed that adults of this species care whether vocal exchanges respect the call-matching, while juveniles do not. A report suggested that a social rule guides western lowland gorillas' contact (grunt) call exchanges, i.e. interlocutors are preferentially selected among group members close in age (Lemasson et al. 2018). A playback experiment broadcasting artificial vocal exchanges respecting or not this social rule showed that older subjects, more experienced, paid more attention to exchanges respecting the age proximity expectation than younger ones (Pougnault et al. 2020).

## 4.4 Conclusion

During recent decades, research focusing on the existence and the developmental acquisition of rules governing nonhuman primates' conversation-like vocal interactions has enlarged perspectives. Current findings bring to light nonhuman primates' ability to take turns during alternated vocal interactions. They select their vocal partner and respect a common response delay to avoid overlap between turns, a delay which seems culturally and socially influenced as in humans. Moreover, nonhuman primates demonstrate vocal elaboration and persistence capacities so as to elicit or to perpetuate a vocal interaction. None of these abilities seem to appear spontaneously but are acquired through social experience during all their ontogeny. Additionally, recent neurobiological arguments support the idea that vocal production during vocal interactions is mediated by neuro-cortical control. Authors suggest volitional control of vocal utterances during social interactions. Human conversation, and more precisely the rules governing interactions, could be homologies (traits shared with other primates) rather than analogies (parallel evolution) (Henry et al. 2015). In any event, these abilities could have been an important step in the evolution of primates' vocal communication.

## References

- Aitken PG (1981) Cortical control of conditioned and spontaneous vocal behavior in rhesus monkeys. *Brain Lang* 13(1):171–184. [https://doi.org/10.1016/0093-934X\(81\)90137-1](https://doi.org/10.1016/0093-934X(81)90137-1)
- Arlet M, Jubin R, Masataka N, Lemasson A (2015) Grooming-at-a-distance by exchanging calls in non-human primates. *Biol Lett* 11(10):20150711–20150714. <https://doi.org/10.1098/rsbl.2015.0711>
- Babel M (2010) Dialect divergence and convergence in New Zealand English. *Lang Soc* 39(4):437–456. <https://doi.org/10.1017/S0047404510000400>
- Babel M (2012) Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *J Phon* 40:177–189. <https://doi.org/10.1016/j.wocn.2011.09.001>
- Bailey WJ (2003) Insect duets: underlying mechanisms and their evolution. *Physiol Entomol* 28:157–174. <https://doi.org/10.1046/j.1365-3032.2003.00337.x>



- Bascom WR (1942) The principle of seniority in the social structure of the Yoruba. *Am Anthropol* 44:37–46. <https://doi.org/10.1525/aa.1942.44.1.02a00050>
- Beach WA (1990) Searching for universal features of conversation. *Res Lang Soc Interact* 24 (1–4):351–368. <https://doi.org/10.1080/08351819009389347>
- Beebe B, Jaffe J, Feldstein S, Mays K, Alson D (1985) Interpersonal timing: the application of an adult dialogue model to mother-infant vocal and kinesic interactions. In: Field T, Fox N (eds) *Social perception in infants*. Ablex Publishing, Norwood, pp 217–247
- Behr O, Knörnschild M, Von Helversen O (2009) Territorial counter-singing in male sac-winged bats *saccopteryx bilineata*: low-frequency songs trigger a stronger response. *Behav Ecol Sociobiol* 63(3):433–442. <https://doi.org/10.1007/s00265-008-0677-2>
- Biben M, Symmes D, Masataka N (1986) Temporal and structural analysis of affiliative vocal exchanges in squirrel monkeys (*Saimiri sciureus*). *Behaviour* 98(1):259–273
- Bouchet H, Pellier AS, Blois-Heulin C, Lemasson A (2010) Sex differences in the vocal repertoire of adult red-capped mangabeys (*Cercocebus torquatus*): a multi-level acoustic analysis. *Am J Primatol* 72(4):360–375. <https://doi.org/10.1002/ajp.20791>
- Bouchet H, Koda H, Lemasson A (2017) Age-dependent change in attention paid to vocal exchange rules in Japanese macaques. *Anim Behav* 129:81–92. <https://doi.org/10.1016/j.anbehav.2017.05.012>
- Briseno-Jaramillo M, Ramos-Fernández G, Palacios-Romo TM, Sosa-López JR, Lemasson A, Sosa-López JR, Lemasson A (2018) Age and social affinity effects on contact call interactions in free-ranging spider monkeys. *Behav Ecol Sociobiol* 72:192. <https://doi.org/10.1007/s00265-018-2615-2>
- Calame-Griaule G (1965) *Ethnologie et langage. la parole chez les dogon*. Gallimard, Paris
- Candiotti A, Zuberbühler K, Lemasson A (2012) Convergence and divergence in Diana monkey vocalizations. *Biol Lett* 8(3):382–385. <https://doi.org/10.1098/rsbl.2011.1182>
- Carter GG, Skowronski MD, Faure PA, Fenton B (2008) Antiphonal calling allows individual discrimination in white-winged vampire bats. *Anim Behav* 76:1343–1355. <https://doi.org/10.1016/j.anbehav.2008.04.023>
- Caskey M, Stephens B, Tucker R, Vohr B (2014) Adult talk in the NICU with preterm infants and developmental outcomes. *Pediatrics* 133(3):e578–e584. <https://doi.org/10.1542/peds.2013-0104>
- Chartrand TL, van Baaren R (2009) Human mimicry. *Adv Exp Soc Psychol* 41:219–274. [https://doi.org/10.1016/S0065-2601\(08\)00405-X](https://doi.org/10.1016/S0065-2601(08)00405-X)
- Chen HC, Kaplan G, Rogers LJ (2009) Contact calls of common marmosets (*Callithrix jacchus*): influence of age of caller on antiphonal calling and other vocal responses. *Am J Primatol* 71:165–170. <https://doi.org/10.1002/ajp.20636>
- Choi JY, Takahashi DY, Ghazanfar AA (2015) Cooperative vocal control in marmoset monkeys via vocal feedback. *J Neurophysiol* 114:274–283. <https://doi.org/10.1152/jn.00228.2015>
- Chow CP, Mitchell JF, Miller CT (2015) Vocal turn-taking in a non-human primate is learned during ontogeny. *Proc R Soc B Biol Sci* 282:20150069. <https://doi.org/10.1098/rspb.2015.0069>
- Clark AP, Wrangham RW (1993) Acoustic analysis of wild chimpanzee pant hoots: do Kibale Forest chimpanzees have an acoustically distinct food arrival pant hoot? *Am J Primatol*. <https://doi.org/10.1002/ajp.1350310203>
- Clutton-Brock TH, Albon SD (1979) The roaring of Red Deer and the evolution of honest advertisement. *Behaviour* 69(4):145–170
- Cossette L, Malcuit G, Pomerleau A, Julien D (1986) Temporal structure of maternal language directed at infants 3 months old. *Can J Psychol* 40(4):414–422
- Coudé G, Ferrari PF, Rodà F, Maranesi M, Borelli E, Veroni V, Monti F, Fogassi L (2011) Neurons controlling voluntary vocalization in the macaque ventral premotor cortex. *PLoS One* 6(11):e2682. <https://doi.org/10.1371/journal.pone.0026822>
- Cui X, Bryant DM, Reiss AL (2012) NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *NeuroImage* 59(3):2430–2437. <https://doi.org/10.1016/j.neuroimage.2011.09.003>



- De Marco A, Rebout N, Massiot E, Sanna A, Sterck EHM, Langermans JAM, Cozzolino R, Thierry B, Lemasson A (2019) Differential patterns of vocal similarity in tolerant and intolerant macaques. *Behaviour* 1(aop):1–25. <https://doi.org/10.1163/1568539X-00003562>
- Dediu D, Levinson SC (2013) On the antiquity of language: the reinterpretation of Neandertal linguistic capacities and its consequences. *Front Psychol* 4(397):1–17. <https://doi.org/10.3389/fpsyg.2013.00397>
- Digweed SM, Fedigan LM, Rendall D (2007) Who cares who calls? Selective responses to the lost calls of socially dominant group members in the white-faced capuchin (*Cebus capucinus*). *Am J Primatol* 69:829–835. <https://doi.org/10.1002/ajp.20398>
- Dunbar RIM (1991) Functional significance of social grooming in primates. *Folia Primatol* 57:121–131
- Dunbar RIM (1996) Grooming, gossip and the evolution of language. Faber & Faber, London
- Dunbar RIM (2009) The social brain hypothesis and its implications for social evolution. *Ann Hum Biol* 36(5):562–572. <https://doi.org/10.1080/03014460902960289>
- Dunbar RIM, Marriott A, Duncan NDC (1997) Human conversational behavior. *Hum Nat* 8(3):231–246. <https://doi.org/10.1007/BF02912493>
- Duncan S (1972) Some signals and rules for taking speaking turns in conversations. *Soc Psychol* 23(2):283–292
- Fedurek P, Schel AM, Slocombe KE (2013) The acoustic structure of chimpanzee pant-hooting facilitates chorusing. *Behav Ecol Sociobiol* 67(11):1781–1789. <https://doi.org/10.1007/s00265-013-1585-7>
- Ferrari PF, Gallese V, Rizzolatti G, Fogassi L (2003) Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur J Neurosci* 17(8):1703–1714. <https://doi.org/10.1046/j.1460-9568.2003.02601.x>
- Fogassi L, Ferrari PF (2007) Mirror neurons and the evolution of embodied language. *Curr Dir Psychol Sci* 16(136). <https://doi.org/10.1353/lan.2005.0202>
- Forester DC, Harrison WK (2008) The significance of persistent vocalisation by the spring peeper, *Pseudacris Crucifer* (Anura: *Hylidae*). *Behaviour* 103(1/3):1–15. <https://doi.org/10.1163/156853989x00303>
- Fuentes A (1998) Re-evaluating primate monogamy. *Am Anthropol* 100(4):890–907
- Gallois C, Ogay T, Giles H (2005) Communication accommodation theory a look back and a look ahead. In: Gudykunst B (ed) *Theorizing about intercultural communication*. Sage, Thousand Oaks, pp 121–148
- Geissmann T (2002) Duet-splitting and the evolution of gibbon songs. *Biol Rev Camb Philos Soc* 77:57–76. <https://doi.org/10.1017/S1464793101005826>
- Geissmann T, Orgeldinger M (2000) The relationship between duet songs and pair bonds in siamangs, *Hylobates syndactylus*. *Anim Behav* 60:805–809. <https://doi.org/10.1006/anbe.2000.1540>
- Ghazanfar AA, Smith-Rohrberg D, Pollen AA, Hauser MD (2002) Temporal cues in the antiphonal long-calling behaviour of cottontop tamarins. *Anim Behav* 64:427–438. <https://doi.org/10.1006/anbe.2002.3074>
- Ghiglione R (1986) *L'homme communicant*. (A. Colin, Ed.). Paris, France
- Ginsburg GP, Kilbourne BK (1988) Emergence of vocal alternation in mother-infant interchanges. *J Child Lang* 15:221–235. <https://doi.org/10.1017/S0305000900012344>
- Golinkoff RM (1986) 'I beg your pardon?': The preverbal negotiation of failed messages. *J Child Lang* 21:205–226. <https://doi.org/10.1017/S0305000900006826>
- Golinkoff RM (1993) When is communication a "meeting of minds"? *J Child Lang* 20:199–207
- Goll Y, Demartsev V, Koren L, Geffen E (2017) Male hyraxes increase countersinging as strangers become 'nasty neighbours'. *Anim Behav* 134:9–14. <https://doi.org/10.1016/j.anbehav.2017.10.002>
- Gregory SW (1990) Analysis of fundamental frequency reveals covariation in interview partners' speech. *J Nonverbal Behav* 14(4):237–251. <https://doi.org/10.1007/BF00989318>

- Griffin ZM, Bock K (2000) What the eyes say about speaking. *Psychol Sci* 11(4):274–279. <https://doi.org/10.1111/1467-9280.00255>
- Grubb P, Butynski TM, Oates JF, Bearder SK, Disotell TR, Groves CP, Struhsaker TT (2003) Assessment of the diversity of African Primates. *Int J Primatol* 24(6):1301–1357
- Hage SR, Nieder A (2013) Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nat Commun* 4:2409. <https://doi.org/10.1038/ncomms3409>
- Haimoff EH (1986) Convergence in the duetting of monogamous Old World primates. *J Hum Evol* 15(1):51–59. [https://doi.org/10.1016/S0047-2484\(86\)80065-3](https://doi.org/10.1016/S0047-2484(86)80065-3)
- Hall ET (1959) *The silent language*. Doubleday, New York
- Hartbauer M, Kratzer S, Steiner K, Römer H (2005) Mechanisms for synchrony and alternation in song interactions of the bushcricket *Mecopoda elongata* (Tettigoniidae: *Orthoptera*). *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 191(2):175–188. <https://doi.org/10.1007/s00359-004-0586-4>
- Hauser MD, Marler P (1993) Food-associated calls in rhesus macaques (*Macaca mulatta*): II. Costs and benefits of call production and suppression. *Behav Ecol* 4:206–212
- Hedwig B (2006) Pulses, patterns and paths: neurobiology of acoustic behaviour in crickets. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol*. <https://doi.org/10.1007/s00359-006-0115-8>
- Heldner M, Edlund J (2010) Pauses, gaps and overlaps in conversations. *J Phon* 38(4):555–568. <https://doi.org/10.1016/j.wocn.2010.08.002>
- Henry L, Craig AJFK, Lemasson A, Hausberger M (2015) Social coordination in animal vocal interactions. Is there any evidence of turn-taking? The starling as an animal model. *Front Psychol* 6:1416. <https://doi.org/10.3389/fpsyg.2015.01924>
- Hilbrink EE, Gattis M, Levinson SC (2015) Early developmental changes in the timing of turntaking: a longitudinal study of mother-infant interaction. *Front Psychol* 6:1492. <https://doi.org/10.3389/978-2-88919-825-2>
- Hyman J (2003) Countersinging as a signal of aggression in a territorial songbird. *Anim Behav* 65(6):1179–1185. <https://doi.org/10.1006/anbe.2003.2175>
- Indefrey P, Levelt WJM (2004) The spatial and temporal signatures of word production components. *Cognition* 92(1–2):101–144. <https://doi.org/10.1016/j.cognition.2002.06.001>
- Jaffe J, Beebe B, Feldstein S, Crown CL, Jasnow MD (2001) Rhythms of dialogue in infancy: coordinated timing in development. *Monogr Soc Res Child Dev* 66(2): 1–132. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11428150>
- Jiang J, Dai B, Peng D, Zhu C, Liu L, Lu C (2012) Neural synchronization during face-to-face communication. *J Neurosci* 32(45):16064–16069. <https://doi.org/10.1523/jneurosci.2926-12.2012>
- Jordan K, Weiss D, Hauser M, McMurray B (2004) Antiphonal responses to loud contact calls produced by *Saguinus oedipus*. *Int J Primatol* 25(2):465–475
- Jürgens U (1995) Neuronal control of vocal production in non-human and human Primates. In: Zimmermann JU, Newman JD (eds) *Current topics in primate vocal communication*. Springer, Boston, pp 199–206. [https://doi.org/10.1007/978-1-4757-9930-9\\_10](https://doi.org/10.1007/978-1-4757-9930-9_10)
- Jürgens U, Ploog D (1970) Cerebral representation of vocalization in the squirrel monkey. *Exp Brain Res* 10(5):532–554. <https://doi.org/10.1007/BF00234269>
- Jürgens U, Maurus M, Ploog D, Winter P (1967) Vocalization in the squirrel monkey (*Saimiri sciureus*) elicited by brain stimulation. *Exp Brain Res* 4:114–117. <https://doi.org/10.1007/BF00240356>
- Jürgens U, Lu C-L, Quondammatteo F (1996) C-fos expression during vocal mobbing in the new world monkey *Saguinus fuscicollis*. *Eur J Neurosci* 8(1):2–10. <https://doi.org/10.1111/j.1460-9568.1996.tb01162.x>
- Kappeler PM, Van Schaik CP (2002) Evolution of primate social systems. *Int J Primatol* 23(4):707–740
- Keller H, Lohaus A, Völker S, Cappenberg M, Chasiotis A (1999) Temporal contingency as an independent component of parenting behavior. *Child Dev* 70(2):474–485. <https://doi.org/10.1111/1467-8624.00034>
- Kerbat-Orecchioni C (1990) *Les interactions verbales*. Armand Collin, Paris

- Koda H (2004) Flexibility and context-sensitivity during the vocal exchange of coo calls in wild Japanese macaques (*Macaca fuscata yakui*). *Behaviour* 141(10):1279–1296. <https://doi.org/10.1163/1568539042729685>
- Kohler E, Keysers C, Umiltà MA, Fogassi L, Gallese V, Rizzolatti G (2002) Hearing sounds, understanding actions: action representation in Mirror neurons. *Science* 297:846–848
- Kurtić E, Brown GJ, Wells B (2013) Resources for turn competition in overlapping talk. *Speech Comm* 55:1–23. <https://doi.org/10.1016/j.specom.2012.10.002>
- Laporte MNC, Zuberbühler K (2010) Vocal greeting behaviour in wild chimpanzee females. *Anim Behav* 80(3):467–473. <https://doi.org/10.1016/j.anbehav.2010.06.005>
- Leavens DA, Hostetter AB, Wesley MJ, Hopkins WD (2004) Tactical use of unimodal and bimodal communication by chimpanzees, Pan troglodytes. *Anim Behav* 67(3):467–476. <https://doi.org/10.1016/j.anbehav.2003.04.007>
- Leavens DA, Russell JL, Hopkins WD (2009) Multimodal communication by captive chimpanzees (Pan troglodytes). *Anim Cogn* 13(1):33–40. <https://doi.org/10.1007/s10071-009-0242-z>
- Lemasson A (2011) What can forest guenons « tell » us about the origin of language? In: Vilain A, Schwartz JL, Vauclair JC (eds) Primate communication and human language: vocalisation, gestures, imitation and deixis in humans and non-humans. John Benjamins Publishing, Amsterdam, pp 39–70
- Lemasson A, Hausberger M (2004) Patterns of vocal sharing and social dynamics in a captive group of Campbell's monkeys (*Cercopithecus campbelli campbelli*). *J Comp Psychol* 118(3):347–359. <https://doi.org/10.1037/0735-7036.118.3.347>
- Lemasson A, Hausberger M (2011) Acoustic variability and social significance of calls in female Campbell's monkeys (*Cercopithecus campbelli campbelli*). *J Acoust Soc Am* 129(5):3341–3352. <https://doi.org/10.1121/1.3569704>
- Lemasson A, Gautier JP, Hausberger M (2003) Vocal similarities and social bonds Campbell's monkey (*Cercopithecus campbelli*). *C R Biol* 326(12):1185–1193. <https://doi.org/10.1016/j.crvi.2003.10.005>
- Lemasson A, Hausberger M, Zuberbühler K (2005) Socially meaningful vocal plasticity in adult Campbell's monkeys (*Cercopithecus campbelli*). *J Comp Psychol* 119:220–229. <https://doi.org/10.1037/0735-7036.119.2.220>
- Lemasson A, Blois-Heulin C, Jubin R, Hausberger M (2006) Female social relationships in a captive group of Campbell's monkeys (*Cercopithecus campbelli campbelli*). *Am J Primatol* 68:1161–1170. <https://doi.org/10.1002/ajp.20315>
- Lemasson A, Gandon E, Hausberger M (2010) Attention to elders' voice in non-human primates. *Biol Lett* 6(3):325–328. <https://doi.org/10.1098/rsbl.2009.0875>
- Lemasson A, Glas L, Barbu S, Lacroix A, Guilloux M, Remeuf K, Koda H (2011a) Youngsters do not pay attention to conversational rules: is this so for nonhuman primates? *Sci Rep* 1:12–15. <https://doi.org/10.1038/srep00022>
- Lemasson A, Ouattara K, Petit EJ, Zuberbühler K (2011b) Social learning of vocal structure in a nonhuman primate? *BMC Evol Biol* 11(362):1–7. <https://doi.org/10.1186/1471-2148-11-362>
- Lemasson A, Guilloux M, Rizaldi, Barbu S, Lacroix A, Koda H (2013a) Age- and sex-dependent contact call usage in Japanese macaques. *Primates* 54:283–291. <https://doi.org/10.1007/s10329-013-0347-5>
- Lemasson A, Ouattara K, Zuberbühler K (2013b) Exploring the gaps between primate calls and human language. In: Botha M, Everaert R (eds) The evolutionary emergence of language: evidence and inference. Oxford University Press, Utrecht, pp 181–203
- Lemasson A, Jubin R, Masataka N, Arlet M (2016) Copying hierarchical leaders' voices? Acoustic plasticity in female Japanese macaques. *Sci Rep* 6(21289). <https://doi.org/10.1038/srep21289>
- Lemasson A, Pereira H, Levréro F (2018) Social basis of vocal interactions in Western lowland gorillas (*Gorilla g. gorilla*). *J Comp Psychol* 132(2):141–151. <https://doi.org/10.1037/com0000105>
- Levinson SC (1983) *Pragmatics*. Cambridge University Press, Cambridge

- Levinson SC (2016) Turn-taking in human communication – origins and implications for language processing. *Trends Cogn Sci* 20(1):6–14. <https://doi.org/10.1016/j.tics.2015.10.010>
- Levinson SC, Holler J (2014) The origin of human multi-modal communication. *Philos Trans R Soc Lond B Biol Sci* 369(1651). <https://doi.org/10.1098/rstb.2013.0302>
- Lévréro F, Touitou S, Frédet J, Nairaud B, Guéry J-P, Lemasson A (2019) Social bonding drives vocal exchanges in bonobos. *Sci Rep* 9:711. <https://doi.org/10.1038/s41598-018-36024-9>
- Lindsay L, Gambi C, Rabagliati H (2019) Preschoolers optimize the timing of their conversational turns through flexible coordination of language comprehension and production. *Psychol Sci* 30(4):504–515
- MacKinnon J, MacKinnon K (1980) The behavior of wild spectral tarsiers. *Int J Primatol* 1(4):361–379. <https://doi.org/10.1007/BF02692280>
- MacLarnon A, Hewitt G (2004) Increased breathing control: another factor in the evolution of human language. *Evol Anthropol* 13:181–197. <https://doi.org/10.1002/evan.20032>
- Masataka N (2003) The onset of language. In: *The onset of language*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511489754>
- Masataka N, Biben M (1987) Temporal rules regulating affiliative vocal exchanges of squirrel monkeys. *Behaviour* 101:311–319
- McComb K, Semple S (2005) Coevolution of vocal communication and sociality in primates. *Biol Lett* 1(4):381–385. <https://doi.org/10.1098/rsbl.2005.0366>
- McGarva AR, Warner RM (2003) Attraction and social coordination: mutual entrainment of vocal activity rhythms. *J Psycholinguist Res* 32(3):335–354. <https://doi.org/10.1023/A:1023547703110>
- Mennill DJ, Boag PT, Ratcliffe LM (2003) The reproductive choices of eavesdropping female black-capped chickadees, *Poecile atricapillus*. *Naturwissenschaften* 90(12):577–582. <https://doi.org/10.1007/s00114-003-0479-3>
- Miller CT, Wang X (2006) Sensory-motor interactions modulate a primate vocal behavior: antiphonal calling in common marmosets. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 192:27–38. <https://doi.org/10.1007/s00359-005-0043-z>
- Miller CT, DiMauro A, Pistorio A, Hendry S, Wang X (2010) Vocalization induced CFos expression in marmoset cortex. *Front Integr Neurosci* 4(128). <https://doi.org/10.3389/fnint.2010.00128>
- Mitani JC, Nishida T (1993) Contexts and social correlates of long-distance calling by male chimpanzees. *Anim Behav* 45(4):735–746. <https://doi.org/10.1006/anbe.1993.1088>
- Nguyen N, Delvaux V (2015) Role of imitation in the emergence of phonological systems. *J Phon* 53:46–54. <https://doi.org/10.1016/j.wocn.2015.08.004>
- Nishitani N, Schürmann M, Amunts K, Hari R (2005) Broca's region: from action to language. *Physiology* 20(1):60–69. <https://doi.org/10.1152/physiol.00043.2004>
- Nummela SU, Jovanovic V, de la Mothe L, Miller CT (2017) Social context-dependent activity in marmoset frontal cortex populations during natural conversations. *J Neurosci* 37(29):7036–7047. <https://doi.org/10.1523/JNEUROSCI.0702-17.2017>
- Okobi DEJ, Banerjee A, Matheson AMM, Phelps SM, Long MA (2019) Motor cortical control of vocal interaction in neotropical singing mice. *Science* 363(6430):979–983. <https://doi.org/10.1126/science.aau1758>
- Ouattara K, Lemasson A, Zuberbühler K (2009a) Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proc Natl Acad Sci* 106:22026–22031. <https://doi.org/10.1073/pnas.0908118106>
- Ouattara K, Lemasson A, Zuberbühler K (2009b) Campbell's monkeys use affixation to alter call meaning. *PLoS One* 4:e7808. <https://doi.org/10.1371/journal.pone.0007808>
- Papoušek M, Papoušek H (1989) Forms and functions of vocal matching in interactions between mothers and their precanonical infants. *First Lang* 9(6):137–157. <https://doi.org/10.1177/014272378900900603>
- Pardo JS, Gibbons R, Suppes A, Krauss RM (2012) Phonetic convergence in college roommates. *J Phon* 40:190–197. <https://doi.org/10.1016/j.wocn.2011.10.001>

- Pika S, Wilkinson R, Kendrick KH, Vernes SC (2018) Taking turns: bridging the gap between human and animal communication. *Proc R Soc B* 285(1880):1–9. <https://doi.org/10.1098/rspb.2018.0598>
- Pollick AS, Gouzoules H, De Waal FBM (2005) Audience effects on food calls in captive brown capuchin monkeys, *Cebus apella*. *Anim Behav* 70(6):1273–1281. <https://doi.org/10.1016/j.anbehav.2005.03.007>
- Pougnault L, Levréro F, Mulot B, Lemasson A (2020) Breaking conversational rules matters to captive gorillas: A playback experiment. *Sci Rep* in press
- Reby D, McComb K (2003) Vocal communication and reproduction in deer. *Adv Study Behav* 33:231–264. [https://doi.org/10.1016/S0065-3454\(03\)33005-0](https://doi.org/10.1016/S0065-3454(03)33005-0)
- Reissland N, Stephenson T (1999) Turn-taking in early vocal interaction: A comparison of premature and term infants' vocal interaction with their mothers. *Child Care Health Dev* 25(6):447–456. <https://doi.org/10.1046/j.1365-2214.1999.00109.x>
- Rizzolatti G, Arbib MA (1998) Language within our grasp. *Trends Neurosci* 21(5):188–194
- Roberts F, Margutti P, Takano S (2011) Judgments concerning the valence of inter-turn silence across speakers of American English, Italian, and Japanese. *Discourse Process* 48:331–354. <https://doi.org/10.1080/0163853X.2011.558002>
- Sacks H, Schegloff EA, Jefferson G (1974) A simplest systematics for the organization of turn-taking for conversation. *Language* 50:696–735. <https://doi.org/10.2307/412243>
- Schaffer R (1977) Study on mother-infant interaction. Academic Press, New York
- Schamberg I, Cheney DL, Clay Z, Hohmann G, Seyfarth RM (2016) Call combinations, vocal exchanges and interparty movement in wild bonobos. *Anim Behav* 122:109–116. <https://doi.org/10.1016/j.anbehav.2016.10.003>
- Schegloff EA (2002) Overlapping talk and the organization of turn-taking for conversation. *Lang Soc* 29(01):1–63. <https://doi.org/10.1017/s0047404500001019>
- Schwartz JJ (1987) The function of call alternation in anuran amphibians: a test of three hypotheses. *Evolution* 41(3):461. <https://doi.org/10.2307/2409249>
- Sewall KB, Young AM, Wright TF (2016) Social calls provide novel insights into the evolution of vocal learning. *Anim Behav* 120:163–172. <https://doi.org/10.1016/j.anbehav.2016.07.031>
- Silk JB (2002) Grunts, Grneys, and good intentions: the origins of strategic commitment in nonhuman primates. In: Nesse R (ed) *Commitment: evolutionary perspectives*. Russell Sage Press, New York, pp 138–157
- Simões CS, Vianney PVR, de Moura MM, Freire MAM, Mello LE, Samesshima K, Araújo J, Nicoletis M, Ribeiro S (2010) Activation of frontal neocortical areas by vocal production in marmosets. *Front Integr Neurosci* 4:1–12. <https://doi.org/10.3389/fnint.2010.00123>
- Slocombe KE, Kaller T, Turman L, Townsend SW, Papworth S, Squibbs P, Zuberbühler K (2010) Production of food-associated calls in wild male chimpanzees is dependent on the composition of the audience. *Behav Ecol Sociobiol* 64(12):1959–1966. <https://doi.org/10.1007/s00265-010-1006-0>
- Snowdon CT, Cleveland J (1984) “Conversations” among pygmy marmosets. *Am J Primatol* 7(1):15–20. <https://doi.org/10.1002/ajp.1350070104>
- Snowdon CT, Elowson AM (1999) Pygmy marmosets modify call structure when paired. *Ethology* 105(10):893–908. <https://doi.org/10.1046/j.1439-0310.1999.00483.x>
- Soltis J, Bernhards D, Donkin H, Newman JD (2002) Squirrel monkey chuck call: vocal response to playback chucks based on acoustic structure and affiliative relationship with the caller. *Am J Primatol* 7:119–130. <https://doi.org/10.1002/ajp.10039>
- Stephens GJ, Silbert LJ, Hasson U (2010) Speaker-listener neural coupling underlies successful communication. *Proc Natl Acad Sci* 107(32):14425–14430. <https://doi.org/10.1073/pnas.1008662107>
- Stern DN, Gibbon J (1979) Temporal expectancies of social behaviours in mother-infant play. In: Thoman EB (ed) *Origins of the infant's social responsiveness*. Lawrence Erlbaum Associates, Hillsdale, pp 409–429

- Stern DN, Jaffe J, Beebe B, Bennett SL (1975) Vocalizing in unison and in alternation: two modes of communication within the mother–infant dyad. *Ann N Y Acad Sci* 263(1):89–100. <https://doi.org/10.1111/j.1749-6632.1975.tb41574.x>
- Stern DN, Beebe B, Jaffe J, Bennett SL (1977) The infant's stimulus world during social interaction: a study of caregiver behaviors with particular reference to repetition and timing. In: Schaffer HR (ed) *Studies in mother–infant interaction*. Academic Press, New York, pp 177–202
- Stivers T, Enfield NJ, Brown P, Englert C, Hayashi M, Heinemann T, Hoymann G, Rossano F, de Ruiter J, Levinson SC (2009) Universals and cultural variation in turn-taking in conversation. *Proc Natl Acad Sci* 106(26):10587–10592. <https://doi.org/10.1073/pnas.0903616106>
- Sugiura H (1993) Temporal and acoustic correlates in vocal exchange of coo calls in Japanese macaques. *Behaviour* 124:3–4. <https://doi.org/10.1163/156853993X00588>
- Sugiura H (1998) Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Anim Behav* 55:673–687. <https://doi.org/10.1006/ambe.1997.0602>
- Sugiura H (2007) Adjustment of temporal call usage during vocal exchange coo calls in Japanese macaques. *Ethology* 113:528–533. <https://doi.org/10.1111/j.1439-0310.2007.01349.x>
- Sugiura H, Masataka N (1995) Temporal and acoustic flexibility in vocal exchanges of coo calls in Japanese macaques (*Macaca fuscata*). In: Zimmermann E, Newman JD, Jürgens U (eds) *Current topics in primate vocal communication*. Plenum Press, New York, pp 121–140. <https://doi.org/10.1007/978-1-4757-9930-9>
- Sutton D, Larson C, Lindeman RC (1974) Neocortical and limbic lesion effects on primate phonation. *Brain Res* 71:61–75
- Takahashi DY, Narayanan DZ, Ghazanfar AA (2013) Coupled oscillator dynamics of vocal turn-taking in monkeys. *Curr Biol* 23:2162–2168. <https://doi.org/10.1016/j.cub.2013.09.005>
- Takahashi DY, Fenley AR, Ghazanfar AA (2016) Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. *Philos Trans R Soc B Biol Sci* 371(1693):20150370. <https://doi.org/10.1098/rstb.2015.0370>
- Tanaka T, Sugiura H, Masataka N (2006) Cross-sectional and longitudinal studies of the development of group differences in acoustic features of coo calls in two groups of Japanese macaques. *Ethology* 112(1):7–21. <https://doi.org/10.1111/j.1439-0310.2006.01103.x>
- Ter Maat M, Truong KP, Heylen D (2010) How turn-taking strategies influence users' impressions of an agent. In: *Intelligent virtual agents: proceedings of the 10th international conference, IVA*, pp 441–453. [https://doi.org/10.1007/978-3-642-15892-6\\_48](https://doi.org/10.1007/978-3-642-15892-6_48)
- Thierry B (2007) Unity in diversity: lessons from macaque societies. *Evol Anthropol* 16(6):224–238. <https://doi.org/10.1002/evan.20147>
- Thierry B, Iwaniuk N, Pellis SM (2000) The influence of phylogeny on the social behaviour of macaques (Primates: *Cercopithecidae*, genus *Macaca*). *Ethology* 106:713–728
- Toarmino CR, Wong L, Miller CT (2017) Audience affects decision-making in a marmoset communication network. *Biol Lett* 13(1):20160934. <https://doi.org/10.1098/rsbl.2016.0934>
- Trudgill P (2008) Colonial dialect contact in the history of European languages: on the irrelevance of identity to new-dialect formation. *Lang Soc* 37(2):241–254. <https://doi.org/10.1017/S0047404508080287>
- Vehrencamp SL, Ellis JM, Cropp BF, Koltz JM (2014) Negotiation of territorial boundaries in a songbird. *Behav Ecol* 25(6):1436–1450. <https://doi.org/10.1093/beheco/aru135>
- Wilson M, Wilson TP (2005) An oscillator model of the timing of turn-taking. *Psychon Bull Rev* 12(6):957–968. <https://doi.org/10.3758/BF03206432>
- Yamaguchi C, Izumi A, Nakamura K (2009) Temporal rules in vocal exchanges of phee and trills in common marmosets (*Callithrix jacchus*). *Am J Primatol* 71:617–622. <https://doi.org/10.1002/ajp.20697>
- Yoshida S, Okanoya K (2005) Evolution of turn-taking: a bio-cognitive perspective. *Cogn Stud* 12(3):153–165
- Zelick R, Narins PM (1985) Characterization of the advertisement call oscillator in the frog *Eleutherodactylus coqui*. *J Comp Physiol A* 156(2):223–229. <https://doi.org/10.1007/BF00610865>



## 4. Problématique et objectifs de la recherche

### 4.1. Problématique

Dans la revue de littérature précédemment développée (**article 1**), il ressort que les échanges vocaux organisés et les règles temporelles et sociales qui les régissent ont fait l'objet de beaucoup d'investigations chez les singes non-Hominidés, révélant de nombreux parallèles avec les conversations humaines. La question se pose alors de savoir si nous avons à faire ici à des comportements analogues issus d'une convergence évolutive entre primates humains et non-humains guidée par les exigences de la vie sociale, ou si nous sommes face à une caractéristique ancestrale avec des règles « conversationnelles » détectables à tous les niveaux de l'échelle phylogénétique des primates. Pourtant, sans qu'il n'y ait d'explications logiques, très peu d'études ont été réalisées chez les grands singes. A l'amorce de ce projet de recherche doctorale, une seule étude menée par Arcadi (2000) chez les mâles chimpanzés de la communauté de Kibale (Ouganda) mettait en lumière l'absence d'échanges vocaux organisés chez cette espèce, au profit de vocalisations isolées ou de *chorus*. Néanmoins, deux articles en cours de publication à cette période, observaient l'existence d'échanges vocaux socialement et temporellement organisés chez les gorilles des plaines de l'Ouest (*Gorilla gorilla gorilla* : Lemasson *et al.* 2018) et les bonobos (Levréro *et al.* 2019). Ces études ont montré un évitement actif de la coupure de parole chez ces deux espèces ainsi qu'un choix des interlocuteurs fondé sur la proximité d'âge chez les gorilles et l'affinité sociale chez les bonobos. Malgré ces résultats encourageants, le manque de données empêchait de pouvoir dresser le moindre scénario évolutif sur les origines phylogénétiques de cette capacité langagière humaine et des règles sous-jacentes.

De ce fait, nous avons tenté de compléter les connaissances parcellaires actuelles en cherchant à mettre en lumière l'existence d'échanges vocaux organisés chez quatre espèces de grands singes que sont les bonobos, les chimpanzés, les gorilles des plaines de l'Ouest et les orangs-outans de Bornéo. Pour cela, nous avons adopté une approche comparative, mêlant les observations aux expérimentations, afin de mieux comprendre l'influence des différents systèmes sociaux (*i.e.* structure et organisation sociale des groupes) sur les convergences et divergences entre les possibles règles sociales et temporelles régissant les échanges vocaux de ces grands singes ainsi que de leur pertinence sociale.



## 4.2. Choix des modèles d'études

Les grands singes constituent une branche particulièrement intéressante de la phylogénie des primates par plusieurs aspects, notamment de par leur proximité taxonomique avec l'espèce humaine. Tous les grands singes, incluant l'humain, font partie du clade des Hominidés. Les investigations en phylogénétique moléculaire permettent de stipuler que les membres du genre *Pongo* sont les premiers à avoir divergés de la lignée ayant donné naissance au genre *Homo*, et ce, il y a entre 9 et 18 millions d'années (Castellano et Munch 2020; Hobolth *et al.* 2011; Upham *et al.* 2019). Puis une seconde divergence, estimée à une date comprise entre 8 et 11 millions d'années (*e.g.* Castellano et Munch 2020; Scally *et al.* 2012; Upham *et al.* 2019), aurait abouti au genre *Gorilla*. Enfin, la séparation des lignées *Pan* et *Homo* aurait eu lieu il y a 4 à 9 millions d'années (*e.g.* Bailey *et al.* 1992; Brunet *et al.* 2002; Chen et Li 2001; Patterson *et al.* 2006; Prüfer *et al.* 2012; Ruvolo *et al.* 1991; Upham *et al.* 2019), positionnant ainsi les chimpanzés et bonobos comme nos plus proches cousins vivants.

De plus, la très grande diversité de systèmes sociaux, aussi bien au niveau de la structure sociale (*i.e.* solitaire – orang-outans : Delgado et van Schaik 2000, uni-mâle/multi-femelles – gorilles : Williamson et Butynski 2013, multi-mâles/multi-femelles – bonobos et chimpanzés : Kano 1982; van Lawick-Goodall 1968)) que de l'organisation sociale (*e.g.* dispersion des individus mâles et/ou femelles à l'approche de la maturité sexuelle : *e.g.* Morrogh-Bernard *et al.* 2011; Stokes *et al.* 2003 ; dominance des mâles ou des femelles adultes selon les espèces : *e.g.* Surbeck *et al.* 2017; van Hooft et van Schaik 1994 ; degré de tolérance entre congénères : *e.g.* Kano 1992; Lehmann et Boesch 2008), font de ces espèces de très bons modèles d'études pour la recherche comparative avec l'humain. Nous nous attendions donc à trouver une grande diversité d'utilisation des vocalisations chez ces espèces qui présentent une grande diversité de systèmes sociaux, et ainsi proposer des scénarios possibles de co-évolution sociale-vocale.

## 4.3. Objectifs de l'étude

Ce projet de recherche doctorale avait pour objectifs de décrire les éventuels échanges vocaux organisés et les règles sous-jacentes chez les grands singes non-humains. Nous discuterons les résultats obtenus sous l'angle des quatre questions définies par Tinbergen en 1963 qui interrogent les causes (*i.e.* mécanismes) d'un comportement, ses fonctions, son histoire développementale (*i.e.* ontogénèse) et son histoire évolutive (*i.e.* phylogénèse). Dans cette optique, nous proposons deux axes d'études :

**Axe 1 :** Dans un premier temps, nous avons testé la pertinence sociale des possibles règles « conversationnelles » récemment mises en évidence chez les gorilles des plaines de l'Ouest par Lemasson et collaborateurs (2018). Pour cela, une approche expérimentale a été adoptée en vue d'étudier le comportement des individus, adultes comme juvéniles, suite à la présentation d'échanges vocaux reconstitués respectant ou non les règles temporelles ou sociales (**article 2**). Au regard des résultats obtenus, nous avons discuté des possibles fonctions sociales du respect de ces règles chez cette espèce ainsi que de l'aspect ontogénétique sous-jacent à leur perception.

**Axe 2 :** Dans un second temps, nous avons cherché à confirmer ou non l'absence d'échanges vocaux organisés chez les chimpanzés telle que décrite par Arcadi (2000) suite à son étude sur les mâles de la communauté sauvage de Kibale (Ouganda). Pour cela, une approche mêlant des observations comportementales et des enregistrements vocaux a été menée sur un groupe captif de chimpanzés en considérant les individus des deux sexes et de toutes classes d'âges (**article 3**). Nous avons ainsi voulu mieux comprendre l'organisation temporelle et sociale de leurs productions vocales en considérant leur système social particulier.

**Axe 3 :** Enfin, nous avons adopté une approche comparative de la diversité des règles « conversationnelles » mises en lumière chez les différentes espèces de grands singes (**article 4**). Cela a été rendu possible par plusieurs collaborations internationales, nous donnant accès à des données provenant de populations sauvages de chimpanzés et de bonobos, et à nos propres observations menées sur des groupes captifs de chimpanzés et d'orangs-outans de Bornéo. Par cette approche, nous avons évalué les facteurs sociaux à l'origine des convergences et divergences entre les espèces étudiées, et nous avons abordé la question de l'évolution de ces capacités probablement proto-conversationnelles au regard de la complexité sociale apparente.



# Chapitre II

## Méthodologie générale



© Illustration de Julie Terrazoni  
Dans l'*Atlas de Zoologie Poétique* par Emmanuelle Pouydebat,  
aux éditions Arthaud



## 1. Les grands singes non-humains

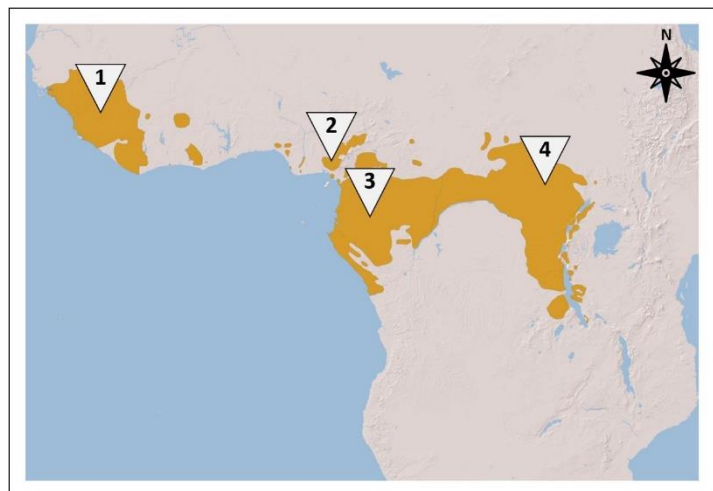
### 1.1. Les chimpanzés

#### 1.1.1. Phylogénèse, répartition géographique et écologie

Les analyses de l'ADN mitochondriale ont permis de révéler l'existence de quatre sous-espèces de chimpanzés génétiquement distinctes (*e.g.* Bradley et Vigilant 2002; Kuhlwilm *et al.* 2016; Oates *et al.* 2007; Prado-Martinez *et al.* 2013) : (i) les chimpanzés de l'Ouest (*Pan troglodytes verus*), (ii) les chimpanzés du Cameroun-Nigeria (*Pan troglodytes eliotti*, anciennement nommés *Pan troglodytes vellerosus*), (iii) les chimpanzés d'Afrique Centrale (*Pan troglodytes troglodytes*) et (iv) les chimpanzés de l'Est (*Pan troglodytes schweinfurthii*). Les sous-espèces d'Afrique de l'Est (*i.e.* *P. t. schweinfurthii*) et d'Afrique Centrale (*i.e.* *P. t. troglodytes*) sont les plus proches génétiquement et constituent un premier clade monophylétique, avec une spéciation tardive estimée à ~100 000 ans (*e.g.* Kuhlwilm *et al.* 2016; Prado-Martinez *et al.* 2013). La sous-espèce *P. t. eliotti*, la plus récemment découverte, et la sous-espèce de l'Ouest (*i.e.* *P. t. verus*) semblent avoir amorcées leur spéciation plus prématurément (~400 000 ans), elles forment un second clade monophylétique (Gonder *et al.* 2011 ; Figure 1).

Figure 1 | Distribution géographique des populations sauvages de chimpanzés. Chaque chiffre correspond à une sous-espèce distincte comme suit : (1) *P. t. verus*, (2) *P. t. eliotti*, (3) *P. t. troglodytes*, (4) *P. t. schweinfurthii*

(Source : IUCN Red List of Threatened Species)



Ces quatre sous-espèces occupent une grande diversité d'habitats, mais se rencontrent particulièrement en forêt tropicale humide primaire et secondaire (*e.g.* Goodall 1986; Nishida 1968; Nowak 1999; Oates *et al.* 2007; van Lawick-Goodall 1968). Ils sont aussi bien terrestres qu'arboricoles, et le temps passé à terre varie en fonction du site d'étude, de l'âge et du sexe des individus (Doran 1996). Les chimpanzés se nourrissent de fruits, de plantes herbacées, de

graines et de protéines animales (*e.g.* œufs, insectes et vertébrés : Goodall 1986; van Lawick-Goodall 1968). Ils prédatent aussi des vertébrés de taille moyenne (*e.g.* cochon sauvage *Potamocheirus larvatus*, et jeune colobe *Colobinae*) prélevés de façon opportuniste, bien que des chasses collaboratives entre mâles aient déjà été référencées (Boesch, 1994; Boesch, 2002; Boesch et Boesch, 1989; Mitani et Watts, 2001).

### 1.1.2. Système social et territorialité

A l'état sauvage, les chimpanzés évoluent dans des communautés multi-mâles/multi-femelles de 20 à 120 individus caractérisées par des phénomènes de fission-fusion (Boesch et Boesch-Achermann 2000; Clark et Wrangham 1994; Goodall 1986; Mitani *et al.* 2002; Nishida 1990; Reynolds 2005; van Lawick-Goodall 1968). En effet, au fil des heures, des jours ou des semaines, des individus de la communauté se regroupent (phénomène de fusion) ou se séparent (phénomène de fission : Clark et Wrangham 1994; Kummer 1971). La taille de ces sous-groupes varie de un à 50 membres (Müller et Mitani 2005; Newton-Fisher 2002ab) et dépend, entre autre, de la distribution des ressources alimentaires (Aureli *et al.* 2008; Wrangham *et al.* 1993).

Les chimpanzés sont une espèce dite mâle-philopatrick, c'est-à-dire que les mâles restent tout au long de leur vie dans la communauté qui les a vu naître, tandis que les femelles se dispersent et migrent dans les communautés voisines une fois la maturité sexuelle atteinte (Boesch et Boesch-Achermann 2000; Mitani *et al.* 2002; Nishida 1969; Reynolds 2005; Wrangham et Smuts 1980). Cette propension que les mâles ont à évoluer toute leur vie dans la même communauté favorise la création de liens sociaux stables et solides entre eux. Ces relations peuvent se maintenir sur des périodes de plus de dix ans (Gilby et Wrangham 2008; Mitani 2009a). Les chimpanzés défendent ardemment leur territoire vis-à-vis des communautés voisines rivales (Nishida, 1979; Goodall, 1986; Boesch et Boesch-Achermann, 2000; Lehmann et Boesch, 2003), donnant lieu à des conflits pouvant leur être fatals (Gómez *et al.* 2016; Mitani 2009b). Les alliances de long terme, associées à des coalitions stratégiques ponctuelles entre mâles, jouent alors un rôle essentiel dans la défense et la conquête de territoires chez cette espèce où la compétition inter-communauté est très vive (van Hooff et van Schaik 1994). En outre, au sein même des communautés, il existe une forte compétition entre mâles pour l'accès à un statut social élevé permettant aux individus de hauts rangs de monopoliser l'accès aux femelles (Boesch et Boesch-Achermann 2000; Müller et Mitani 2005; Newton-Fisher *et al.* 2010; Vigilant *et al.*, 2001). La création d'alliances pour un mâle est indispensable pour établir et maintenir son statut social (Müller et Mitani 2005; Nishida 1983; Nishida et Hosaka 1996).

Cela se confirme par le nombre élevé d'alliances que comptent les individus de haut rang social, en comparaison aux mâles de rang inférieur (Duffy 2006).

Ainsi, les mâles constituent le noyau social du groupe. Leur stature est légèrement plus imposant que celle des femelles (*e.g.* Leigh et Shea 1995), et ils les dominent socialement (De Waal 1982). Les femelles évoluent plus en périphérie de la communauté (De Waal 1982), seules avec leur jeune ou accompagnées de quelques femelles et leurs petits (*e.g.* Arnold et Whiten 2003; Wrangham *et al.* 1992). Bien que les femelles chimpanzés soient généralement considérées comme solitaires, Lehmann et Boesch (2008, 2009a) décrivent une socialité complexe et des liens sociaux solides entre elles. De plus, elles ne sont pas exemptes d'une certaine forme de hiérarchie (Lehmann et Boesch 2009b), notamment pour l'accès à la reproduction (Pusey *et al.* 1997; Murray *et al.* 2006). Enfin, il existe peu de relations entre les individus de sexe opposé hors période de reproduction (Stumpf et Boesch, 2005), à l'exception des interactions mère/enfant (Goodall 1986; van Lawick-Goodall 1968). Des différences inter-communautés marquées sont tout de même à noter (Wakefield, 2002; Williams *et al.* 2002).

### 1.1.3. Communication

La communication des chimpanzés regroupent un ensemble de signaux visuels gestuels et faciaux (Graham *et al.* 2018; Hobaiter et Byrne 2011). Toutefois, dans un environnement particulièrement obstrué par un couvert végétal dense, l'utilisation de vocalisations est un avantage évolutif considérable (Marler et Tenaza 1977; Slocombe et Zuberbühler 2010).

Le répertoire vocal de cette espèce est considéré comme gradé (Marler 1976; Marler et Tenaza 1977; Slocombe et Zuberbühler 2010), compliquant de ce fait sa description. Les auteurs dénombrent entre 13 et une trentaine de vocalisations différentes en fonction des communautés (*e.g.* *P. t. verus* Gombe (Tanzanie) : Goodall 1986; *P. t. schweinfurthii* Budongo-Sonso (Ouganda) : Slocombe et Zuberbühler 2010). L'activité vocale des individus varie selon leur classe d'âge, leur sexe, leur statut hiérarchique, leur communauté et le contexte social dans lequel les individus se trouvent (Clark 1993; Slocombe et Zuberbühler 2010). Ainsi, au sein de la population de Budongo, les mâles adultes sont les individus les plus vocaux quand ils sont isolés des membres de leur communauté, alors que les femelles en période de menstruation sont très actives vocalement en présence du groupe (Slocombe et Zuberbühler 2010). Cependant, dans d'autres communautés, comme à Kibale (*P. t. schweinfurthii*, Ouganda), les mâles de haut rang social sont bien plus actifs vocalement que les femelles adultes ou les mâles de rang hiérarchique inférieur (Clark 1993). L'utilisation de certaines vocalisations est, elle aussi, médiée par le sexe et le rang hiérarchique des individus (Marler et Tenaza 1977).



Depuis quelques années, un certain nombre de capacités socio-cognitives et structurelles a été mis en lumière dans la communication vocale des chimpanzés. Leurs vocalisations peuvent notamment encoder des informations intrinsèques telles que l'âge, le sexe, l'identité de l'émetteur (*e.g.* Fedurek *et al.* 2016), ainsi que des informations plus labiles comme le rang social, l'appartenance à une communauté (*e.g.* Arcadi 1996; Crockford *et al.* 2004; Fedurek *et al.* 2016; Notman et Rendall 2005) ou un état émotionnel particulier (*e.g.* Slocombe et Zuberbühler 2010). Des informations concernant des éléments de leur environnement, telles que la nature de l'item alimentaire rencontré, peuvent également être transmises par les vocalisations (*e.g.* Slocombe et Zuberbühler 2005).

Par ailleurs, les chimpanzés sont en mesure de combiner leurs vocalisations afin d'accroître la diversité de leur répertoire vocale (Crockford et Boesch 2005; Leroux et Townsend 2020). Toutefois, les nombreuses fonctions que peuvent avoir ces combinaisons vocales (*i.e.* altération ou intensification du message initial, signal vocal adressé à une audience précise : Crockford et Boesch 2005) restent à déterminer (Leroux et Townsend 2020).

Des phénomènes de convergence vocale se manifestent également lors de la participation des individus à un même *chorus* permettant d'établir, de maintenir et de renforcer les liens sociaux entre les interlocuteurs (Fedurek *et al.* 2013ab; Mitani et Gros-Louis 1998), mais aussi de témoigner de son appartenance à un groupe (Arcadi 1996). En captivité, des observations rapportent le cas d'une modification active de la structure d'une vocalisation référentielle (*i.e.* alimentaire) d'individus adultes suite à un contact prolongé avec les membres d'un groupe social dont ils n'étaient pas originaires (Watson *et al.* 2015). De même, des capacités d'innovation vocale ont été révélées chez certains individus en vue d'attirer l'attention des soigneurs humains (*Raspberry call* : Hopkins *et al.* 2007).

La présence et la composition de l'audience sont aussi connues pour influencer la production vocale chez cette espèce (*e.g.* screams : Fedurek *et al.* 2015; copulation calls : Townsend *et al.* 2008). Récemment, certains auteurs ont pu prouver que la communication vocale des chimpanzés était emprunte d'intentionnalité (Crockford *et al.* 2012, 2015; Gruber *et Zuberbühler* 2013; Schel *et al.* 2013 ; voir *Introduction*, partie 2.1.1). C'est notamment le cas des vocalisations d'alarme produites par un individu à l'attention d'une audience naïve du danger (Schel *et al.* 2013).

## 1.2. Les bonobos

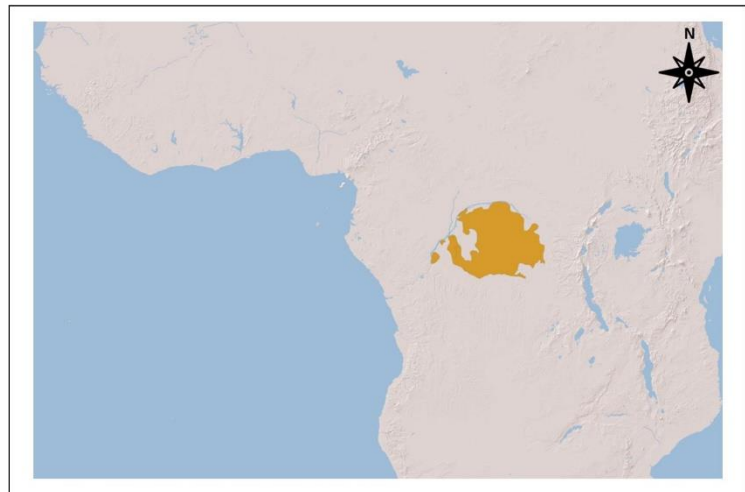
### 1.2.1. Phylogénèse, répartition géographique et écologie

Les bonobos constituent l'autre espèce du genre *Pan* (Kano, 1992). Les analyses génétiques de l'ensemble du génome et de l'ADN mitochondriale des individus ont permis d'estimer que les deux espèces sœurs (*i.e.* *Pan paniscus* et *Pan troglodytes*) ont divergé génétiquement entre ~3 millions d'années et ~1,2 millions d'années (Castellano et Munch 2020; Kuhlwilm *et al.* 2016; Mailund *et al.* 2014; Prado-Martinez *et al.* 2013; Prüfer *et al.* 2012; Upham, Esselstyn, et Jetz 2019).

Cette espèce est endémique de la République Démocratique du Congo (RDC). Les bonobos sont seulement présents dans la forêt tropicale humide primaire et secondaire qui occupe le bassin du Congo, uniquement sur la rive gauche du fleuve Congo. Leur distribution se limite donc par le fleuve Congo à l'Ouest et au Nord, et ses affluents, les rivières Kasai et Sankuru au Sud, et Lualaba à l'Est (Prüfer *et al.* 2012; Prado-Martinez *et al.* 2013; Mailund *et al.* 2014; Kuhlwilm *et al.* 2016; Upham *et al.* 2019; Castellano et Munch 2020; Figure 2).

Figure 2 | Distribution géographique des populations sauvages de bonobos

(Source : IUCN Red List of Threatened Species)



La forêt équatoriale fournit aux bonobos des ressources alimentaires diverses, régulières et abondantes (Idani *et al.* 1994; Kano 1992; Pennec *et al.* 2016; White 1998). Ils s'y déplacent aussi bien au sol que dans les strates arborées (Doran 1996; Kano 1992) et se nourrissent en majorité de fruits et de végétaux qu'ils complètent par des racines, des champignons et diverses sources de protéines animales (*e.g.* œufs, insectes et petits vertébrés : Bermejo *et al.* 1994; Idani *et al.* 1994; Kano 1992; Serckx *et al.* 2015; White 1998). Des petits gibiers (*e.g.* jeunes primates

ou *Cephalophinae*) peuvent également être capturés de façon opportuniste (Fruth et Hohmann 2002; Hohmann et Fruth 2008; Surbeck et Hohmann 2008).

### 1.2.2. Système social et territorialité

Les bonobos, comme les chimpanzés, évoluent dans des communautés multi-mâles/multi-femelles de 30 à 80 animaux (Fruth *et al.* 2016, 2013) soumises à une dynamique de fission-fusion (Furuichi *et al.* 2012; Hashimoto *et al.* 1998; Kano 1992; van Elsacker *et al.* 1995). Les sous-groupes se constituent puis se dissolvent après plusieurs heures, jours ou semaines et se composent de cinq à 25 individus (Idani 1991; Kano 1982; Mulavwa *et al.* 2008; White 1998).

Les bonobos sont une espèce mâle-philopatrique (Furuichi *et al.* 2012; Kano 1982; van Elsacker *et al.* 1995) chez laquelle les femelles constituent le noyau social du groupe. Malgré l'absence de lien de parenté, des coalitions solides et durables se forment entre elles, et de forts liens affiliatifs apparaissent (De Waal 1995a; Furuichi 1989; Parish 1994; Stevens *et al.* 2006; Surbeck *et al.* 2017; Tokuyama et Furuichi 2016). Comparativement, les relations entre mâles sont plus rares et souvent de nature agonistique (Furuichi et Ihobe, 1994; Kano, 1992; Stevens *et al.*, 2006), ce qui contraste profondément avec les relations sociales pouvant exister chez les chimpanzés (Newton-Fisher, 2002ab). Les coalitions entre femelles leur permettent de dominer (ou co-dominer) les mâles, chez lesquels aucune forme de soutien intra-sexe ne semble émerger lors des conflits. Les relations de dominance entre les femelles s'expriment majoritairement en contexte alimentaire (Vervaecke *et al.* 1999).

Les bonobos se caractérisent par leurs comportements socio-sexuels qui se manifestent entre individus de tous âges et de tous sexes. Ces comportements jouent un rôle central dans le maintien et le renforcement des liens sociaux entre les participants, permettant l'apaisement des tensions sociales (*i.e.* consolation) et le règlement des conflits (*i.e.* réconciliation) au sein du groupe (*e.g.* De Waal 1995b, 1997; Furuichi 1989; Hohmann et Fruth 2000; Moscovice *et al.* 2015).

Les bonobos ne défendent pas de frontières territoriales et sont considérés comme plus tolérants et considérablement moins agressifs que les chimpanzés à l'encontre des membres de leur communauté et des communautés voisines (*e.g.* Boesch *et al.* 2002; Furuichi 2011; Wilson *et al.* 2014). En effet, des rencontres entre membres de groupes limitrophes peuvent se produire plusieurs fois par jour (Furuichi 2011; Kano 1992, 1982) et se soldent la plupart du temps par un évitement des individus ou des interactions sociales pacifiques temporaires (Furuichi 2011; Hohmann 2001).

### 1.2.3. Communication

Le répertoire des bonobos comprend des expressions faciales et des gestes communicatifs (Graham *et al.* 2017, 2018), ainsi qu'un ensemble de signaux vocaux (Bermejo et Omedes 1999; Clay *et al.* 2015; De Waal 1988) indispensables dans un environnement où l'accès visuel aux congénères est restreint par le couvert végétal (*e.g.* Gautier 2001; Marler 1965).

Pour l'heure, seules deux études ont tenté de décrire de façon non-exhaustive leur répertoire vocal : l'une en captivité menée par De Waal (1988), et une seconde en milieu sauvage conduite par Bermejo et Omedes (1999). Les auteurs dénombrent entre 12 types vocalisations distinctes en captivité et 15 en milieu sauvage, mais le caractère gradé des signaux vocaux de cette espèce complexifie l'étude de sa communication vocale (Keenan *et al.* 2020). La plupart des investigations menées jusqu'à présent n'ont pas tenu compte de l'influence des facteurs individuels (*i.e.* âge, sexe et rang social) sur le taux de production vocale des bonobos. Il semble toutefois que les femelles adultes ainsi que les mâles sub-adultes soient particulièrement actifs vocalement Cornec et collaborateurs (Soumis).

Par ailleurs, il a été démontré que les vocalisations des bonobos peuvent encoder des messages relatifs à l'identité de l'émetteur (*e.g.* Keenan *et al.* 2020, 2016) ou des informations plus labiles faisant référence à leur état émotionnel (Clay *et al.* 2015). Ils sont aussi capables de transmettre des informations concernant l'environnement extérieur, notamment à l'approche d'items alimentaires (*e.g.* Bermejo et Omedes 1999; Clay et Zuberbühler 2009; De Waal 1988).

Des capacités combinatoires sont également rapportées chez cette espèce (Leroux et Townsend 2020). En contexte alimentaire, l'émetteur peut encoder l'appétence de l'aliment qu'il rencontre par la combinaison de différentes vocalisations (Clay et Zuberbühler 2009). Des expériences de repasse de vocalisations ont démontré que les membres de l'audience étaient en capacité de décoder l'information transmise par ces différentes séquences de vocalisations combinées (Clay et Zuberbühler 2011). La combinaison de différentes vocalisations semble également utilisée pour la coordination socio-spatiale des individus, mais des validations expérimentales sont encore nécessaires (Schamberg *et al.* 2016).

Bien que les vocalisations des primates non-humains semblent majoritairement déterminées génétiquement (*e.g.* Hammerschmidt et Fischer 2008), une certaine flexibilité vocale se manifeste tout de même chez les bonobos par des phénomènes de convergence vocale entre individus ayant de forts liens affinitaires (Levréro *et al.* 2017).

Les bonobos sont aussi sensibles à la présence et à la composition de l'audience. Par exemple, leurs vocalisations de copulations diffèrent en fonction du rang social de l'individu avec lequel ils sont engagés dans un comportement socio-sexuel (Clay *et al.* 2011), mais aussi en fonction de l'audience pouvant percevoir ces vocalisations (Clay et Zuberbühler 2012). Il n'existe pas, pour l'heure, de preuves claires de l'intentionnalité sous-jacente à la communication vocale chez cette espèce. Toutefois, des éléments prometteurs peuvent apparaître lorsqu'un individu quitte son sous-groupe pour en rejoindre un autre. Avant d'entamer son trajet, l'animal produit une séquence vocale et n'engage pas de déplacement tant qu'il n'obtient pas de réponse du sous-groupe d'accueil. Si la réponse se fait attendre, l'individu persiste et réitère son appel (Schamberg *et al.* 2016). Afin de valider avec certitude l'intentionnalité qui sous-tend ce comportement, des expérimentations complémentaires sont indispensables

### 1.3. Les gorilles

#### 1.3.1. Phylogénèse, répartition géographique et écologie

L'analyse génétique de l'ADN mitochondriale, couplée à l'étude des variations craniométriques, ont permis de distinguer deux espèces de gorilles que sont : (i) les gorilles de l'Ouest (*Gorilla gorilla*), et (ii) les gorilles de l'Est (*Gorilla beringei*), dont la divergence génétique est estimée à 3 millions d'années (Antonelli *et al.* 2016; Das *et al.* 2014; Kuhlwilm *et al.* 2016). Chacune de ces deux espèces est composée de deux sous-espèces. Les gorilles de l'Ouest avec (i) les gorilles des plaines de l'Ouest (*Gorilla gorilla gorilla*) et (ii) les gorilles de la rivière Cross (*Gorilla gorilla diehli*), ayant divergés il y a ~68 000 ans (McManus *et al.* 2015). Les gorilles de l'Est avec (i) les gorilles des montagnes (*Gorilla beringei beringei*) et (ii) les gorilles des plaines de l'Est (*Gorilla beringei graueri*) dont la date de divergence est estimée à ~150 000 ans (Prado-Martinez *et al.*, 2013; Xue *et al.*, 2015; Figure 3). Par la suite, nous nous focaliserons sur les gorilles des plaines de l'Ouest (*Gorilla gorilla gorilla*).

Figure 3 | Distribution géographique des populations sauvages de gorilles. Chaque chiffre correspond à une sous-espèce distincte comme suit : (1) *G. g. diehli*, (2) *G. g. gorilla*, (3) *G. b. beringei*, (4) *G. g. graueri*.

(Source : IUCN Red List of Threatened Species)



Les gorilles des plaines de l'Ouest évoluent dans les forêts tropicales humides et dans les plaines inondées (Forcina *et al.* 2019; Gatti *et al.* 2004; Levréro *et al.* 2006; Magliocca et Gautier-Hion 2002; Maisels *et al.* 2018) où ils trouvent une grande partie de leur alimentation composée de feuilles et de plantes herbacées (Magliocca et Gautier-Hion 2002; Tutin *et al.* 1991). Des fruits, de l'écorce et des tiges peuvent venir compléter ce régime alimentaire (Bermejo 2004; Tutin et Fernandez 1993; Tutin *et al.* 1991). Les gorilles sont majoritairement terrestres bien qu'ils aient la capacité de se mouvoir dans les arbres occasionnellement (Ankel-Simons 2007).

### 1.3.2. Système social et territorialité

La dynamique et la structure sociale des gorilles des plaines de l'Ouest sont bien moins documentées que celles des chimpanzés et des bonobos. Les gorilles des plaines de l'Ouest forment des groupes sociaux uni-mâle/multi-femelles stables et cohésifs (Aureli *et al.* 2008) de cinq à dix individus, pouvant atteindre 20 membres et exceptionnellement plus de 30 (Bermejo 2004; Williamson et Butynski 2013). Ces groupes se composent généralement d'un mâle adulte, de trois à quatre femelles adultes accompagnées de leurs petits (Breuer *et al.* 2010; Gatti *et al.* 2004; Parnell 2002). Il est rare d'observer des groupes de plus d'un mâle adulte (Parnell 2002; Tutin 1996). Cette espèce exprime un dimorphisme particulièrement marqué. Les mâles adultes sont deux fois plus imposants que les femelles et développent une musculature très prononcée ainsi qu'une crête sagittale proéminente (*e.g.* Breuer *et al.* 2007; Dixson 1998; Smith et Jungers 1997). Le qualificatif de « dos argenté » attribué aux mâles adultes provient de la fourrure grise pâle qui recouvre leur dos à la maturité sexuelle (Rowe 1996).

Le mâle adulte dominant est l'élément central du groupe. Il coordonne les phases d'activités des individus (*e.g.* déplacements entre les sites d'alimentation, fourrages et repos : Margulis *et al.* 2003; Watts 1996), et endosse le rôle de protecteur du groupe face aux

prédateurs (e.g. léopard *Panthera pardus* : Harcourt et Stewart 2007) et au risque d'infanticide que représentent les mâles concurrents, qui semble toutefois moindre chez cette espèce en comparaison aux gorilles des montagnes (Forcina *et al.* 2019; Gómez *et al.* 2016; Harcourt et Stewart 2007). Il tempore également les relations entre les femelles adultes qui s'en tiennent majoritairement à de la faible compétition voir à de l'indifférence (Hagemann *et al.* 2018; Robbins *et al.* 2004).

Chez les gorilles des plaines de l'Ouest, les mâles, comme les femelles, quittent leur groupe natal à l'approche de la maturité sexuelle. La plupart des mâles sub-adultes rejoindront alors des regroupements temporaires non-reproducteurs (Levréro *et al.* 2006; Parnell 2002). Ces groupes non-reproducteurs sont particulièrement instables étant donné le fort taux de migration auquel ils sont soumis (Levréro *et al.* 2006). Les jeunes femelles, quant à elles, rejoindront un mâle solitaire, un autre groupe reproducteur déjà constitué ou, plus rarement, un groupement temporaire non-reproducteur (Forcina *et al.* 2019; Levréro *et al.* 2006; Stokes *et al.* 2003; Tutin 1996).

A l'inverse des chimpanzés, les territoires de plusieurs groupes peuvent se superposer (Forcina *et al.* 2019). La faible territorialité des gorilles des plaines de l'Ouest peut se justifier par le lien de parenté entre les mâles dominants des groupes limitrophes (Bradley *et al.* 2004). Ainsi, bien que la compétition entre mâles adultes puisse être rude, les conflits se restreignent généralement à des comportements d'intimidation et n'aboutissent que très exceptionnellement à des interactions physiques pouvant être fatales aux protagonistes (Caillaud *et al.* 2008; Gómez *et al.* 2016; Sicotte 1993).

### 1.3.3. Communication

La communication vocale des gorilles est encore peu documentée malgré la littérature croissante sur le sujet. A l'image des membres du genre *Pan*, les gorilles de plaines de l'Ouest communiquent par le biais de signaux gestuels, faciaux (Genty *et al.*, 2009) et vocaux (Salmi, Hammerschmidt, et Doran-Sheehy 2013; Stewart et Harcourt 1994) particulièrement utiles dans un environnement visuellement obstrué (Gautier 2001; Marler 1965).

Là encore, la gradation du répertoire vocal de l'espèce complique la quantification du nombre de types vocalisations (Marler 1976). Les confusions apparaissent majoritairement sur les vocalisations de faible intensité et de basse fréquence, tels que les cris de contact (Hedwig *et al.* 2014; Salmi *et al.* 2013). A l'état sauvage, 17 types de cris différents ont pu être extraits sur la base de caractéristiques acoustiques et contextuelles (Salmi *et al.* 2013). Le mâle adulte a le répertoire vocal le plus important de tous les membres du groupe. C'est aussi l'individu le

plus vocalement actif en comparaison aux femelles adultes et aux individus immatures (Salmi *et al.* 2013).

Chez cette espèce, les vocalisations encodent des informations relatives à l'identité de l'émetteur (Salmi *et al.* 2014). Certains signaux vocaux sont associés à des contextes particuliers tels que la présence de nourriture (Luef *et al.* 2016). Des études complémentaires doivent être conduites pour détecter un possible encodage qualitatif et/ou quantitatif de l'item alimentaire proposé.

Les gorilles des plaines de l'Ouest produisent également des combinaisons vocales (Hedwig *et al.* 2014, 2015a). Ces combinaisons impliquent des vocalisations de contact et semblent être employées principalement à la suite de la vocalisation d'un conspécifique (Hedwig *et al.* 2015a). Bien que des investigations soient encore nécessaires pour déterminer la fonction de ces combinaisons, il semble qu'elles jouent un rôle lors des interactions vocales pacifiques (Hedwig *et al.* 2015a; Leroux et Townsend 2020).

Comme le montre Hedwig et collaborateurs (2015b), une prise en compte de l'effet d'audience se manifeste aussi chez cette espèce. En effet, une augmentation du taux de vocalisations de contact et une modulation de la fréquence fondamentale apparaît en fonction de l'accès visuel aux congénères et de la distance les séparant de l'émetteur. Cette flexibilité vocale qui apparaît suite à un effet d'audience est un premier pas pour l'étude de l'intentionnalité des signaux vocaux. Toutefois, aucune étude n'affirme l'existence d'une intentionnalité sous-tendant à la communication vocale chez les gorilles des plaines de l'Ouest.

## 1.4. Les orangs-outans

### 1.4.1. Phylogénèse, répartition géographique et écologie

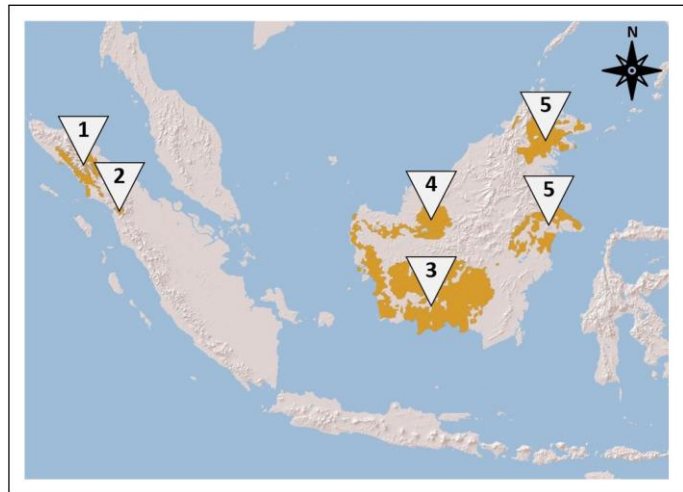
Des distinctions morphologiques et génétiques notables (Groves 2001; Locke *et al.* 2011; Nater *et al.* 2011; Warren *et al.* 2001) ont révélé l'existence de deux espèces d'orangs-outans que sont *Pongo abelii* et *Pongo pygmaeus* (Groves 2001; Mailund *et al.* 2012; Nater *et al.* 2015; Prado-Martinez *et al.* 2013). *P. abelii* évolue sur l'île de Sumatra, tandis que l'île de Bornéo abrite les représentants de l'espèce *P. pygmaeus* (Groves 2001; Wich *et al.* 2008) constitués de trois sous-espèces : (i) *P. p. pygmaeus* au Nord-Ouest de l'île, (ii) *P. p. wurmbii* au centre, et (iii) au Nord-Est, *P. p. morio* (Wich *et al.* 2008; Ancrenaz *et al.* 2018; Figure 4). A l'heure actuelle, aucune sous-espèce n'a été proposée pour *P. abelii* (Greminger *et al.* 2014; Mattle-Greminger 2015; Nater *et al.* 2011). Très récemment, une troisième espèce (*Pongo tapanulienis*) a été décrite par Nater et ses collègues (2017) sur l'île de Sumatra (Figure 4). Par



analyse du génome mitochondriale, les auteurs ont démontré que les représentants de l'espèce *P. tapanuliensis* sont génétiquement plus proches de la lignée de Bornéo, qui auraient divergés il y a ~2,41 millions d'années (entre 1,26 et 3,42 millions d'années : Nater *et al.* 2017). La divergence entre les individus de l'espèce *P. abelii* et le groupe monophylétique constitué des espèces *P. pygmaeus* et *tapanuliensis* aurait eu lieu plus précocement, il y a ~3,97 millions d'années (Nater *et al.* 2017). Par la suite, nous nous intéresserons uniquement à l'espèce évoluant sur l'île de Bornéo.

Figure 4 | Distribution géographique des populations sauvages d'orang-outans. Chaque chiffre correspond à une espèce ou sous-espèce distincte comme suit : (1) *P. abelii*, (2) *P. tapanuliensis*, (3) *P. p. pygmaeus*, (4) *P. p. wurmbii*, (5) *P. p. morio*

(Source : IUCN Red List of Threatened Species)



Les orangs-outans de Bornéo sont presque exclusivement arboricoles (Wich *et al.* 2004, 2009b). Ils occupent la canopée des forêts tropicales humides primaires et des anciennes forêts secondaires de basses altitudes de l'île (Rodman 1988). Ils y trouvent les fruits, les plantes et les graines essentielles à leur régime alimentaire (Kanamori *et al.* 2010; Wich *et al.* 2006). Occasionnellement, ils peuvent se nourrir de viande de petits vertébrés obtenue de façon opportuniste (Sugardjito et Nurhuda 1981; Utami et van Hooff 1997).

#### 1.4.2. Système social et territorialité

Les orangs-outans sont reconnus pour être particulièrement solitaires, aussi bien lors des phases de déplacement que d'alimentation et de repos (Galdikas 1985a; Knott *et al.* 2008; Mackinnon 1974; Rijksen 1978; Rodman 1979). Chez cette espèce femelle-philopatrique (*i.e.* la femelle reste sur son territoire de naissance alors que les jeunes mâles migrent vers d'autres zones géographiques : Delgado et van Schaik 2000; Arora *et al.* 2010; Nater *et al.* 2011; van Noordwijk *et al.* 2012), les interactions sociales représentent en moyenne 1% du budget temps quotidien des individus à l'état sauvage (Knott 1999).

L'unité sociale la plus pérenne est le couple mère-enfant, dont la relation peut s'étaler sur huit années. Ce couple est parfois accompagné d'une des plus jeunes filles de la mère (*e.g.* Galdikas 1985a; Horr 1975; Mackinnon 1974; Mitani 1990; Rijksen 1978). Les territoires des femelles se superposent partiellement (Knott *et al.* 2008) et celui du mâle englobe à lui seul celui de plusieurs femelles, favorisant la probabilité de rencontre en période de reproduction (Delgado et van Schaik 2000). Au cours de cette période, il peut être observé l'association temporaire entre le mâle résident et les femelles adultes (Galdikas 1984; Schürmann 1982), impliquant des comportements affiliatifs et des partages de nourriture (van Schaik *et al.* 2004). L'abondance de nourriture saisonnière dans certaines zones géographiques peut néanmoins favoriser les regroupements d'individus d'âges et de sexes différents (van Schaik *et al.* 2009). Des associations temporaires non-aléatoires entre certains conspécifiques peuvent ainsi émerger (*i.e.* individus qui coordonnent leurs déplacements à une distance maximale de 50 mètres selon Knott et collaborateurs 2008 et van Schaik 1999). La fréquence, la nature et la durée de ces associations sont influencées par l'âge des individus, leur statut reproducteur ainsi que par leur lien de parenté (Knott *et al.* 2008; van Schaik 1999).

Les mâles orangs-outans matures sont connus pour exprimer deux morphologies différentes (Delgado et van Schaik 2000). Les mâles pleinement développés (ou « bi-maturés ») se distinguent des mâles adultes non-développés par l'apparition de caractères sexuels secondaires marqués tels que l'apparition d'excroissances épidermiques au niveau des joues, d'un sac-pharyngé imposant et d'une corpulence près de deux fois supérieure à celle d'une femelle (Rodman et Mitani 1987). Ces mâles bi-maturés tolèrent les mâles adultes non-développés et les sub-adultes, qu'ils dominent aisément (Galdikas 1985bc; Mitani 1985; Rodman 1973). Ils se montrent néanmoins très agressifs à l'encontre d'autres individus ayant atteint le même stade de développement qu'eux (Mackinnon 1974; Mitani 1985), notamment en présence d'une femelle (Galdikas 1985c).

#### 1.4.3. Communication

La communication vocale des orangs-outans de Bornéo est peu répertoriée dans la littérature contemporaine. Cette espèce possède un répertoire communicatif composé de gestes, d'expressions faciales (Liebal *et al.* 2006) et de vocalisations qui sont essentielles dans un environnement au couvert végétal dense (Gautier 2001; Marler 1965). Récemment, Hardus et ses collaborateurs (2009) ont décrit 26 vocalisations distinctes exprimées par les orangs-outans de Bornéo, malgré certaines différences populationnelles (Wich *et al.* 2012). Le répertoire vocal des individus adolescents et adultes est bien plus diversifié que celui des individus immatures,

qui utilisent moitié moins de types de vocalisations (Hardus *et al.* 2009; Spillmann *et al.* 2017). Il existe en revanche peu de différences de taux de production vocale entre les sexes (Delgado et van Schaik 2000; Hardus *et al.* 2009).

Des informations sur l'identité de l'émetteur sont encodées dans les vocalisations de cette espèce (*e.g.* Askew et Morrogh-Bernard 2017). Certains types de vocalisations peuvent également être associés à des contextes particuliers tels que des perturbations biotiques ou abiotiques dans l'environnement proche de l'émetteur (*e.g.* chute d'arbre, présence d'un mâle rival proche : Askew et Morrogh-Bernard 2017; Spillmann *et al.* 2010). Des investigations supplémentaires sont à mener pour déterminer des capacités vocales référentielles chez cette espèce.

Malgré l'existence de vocalisations composées de séquences d'unités sonores combinées chez les orangs-outans de Bornéo (Spillmann *et al.* 2010), il est crucial de poursuivre les études observationnelles et expérimentales afin de décrire d'éventuelles combinaisons vocales chez ces animaux (Leroux et Townsend 2020).

La production de certaines vocalisations semble soumise à un apprentissage social, comme il a été démontré dans certaines populations sauvages (Wich *et al.* 2012). En captivité, il a également pu être démontré que certains individus pouvaient modifier leurs signaux acoustiques et converger ainsi avec les sons (*i.e.* sifflements) produits par l'expérimentateur, témoignant d'une flexibilité vocale indéniable (Lameira *et al.* 2013; Wich *et al.* 2009a).

L'intentionnalité sous-jacente aux productions vocales des orangs-outans reste encore à démontrer.

## 2. Sujets d'étude et conditions de vie

L'étude des comportements en captivité, tout comme en milieu naturel, porte son lot d'avantages et d'inconvénients dû à l'environnement de vie des modèles d'études. Dans un premier temps, les contraintes inhérentes à la captivité réduisent le nombre d'animaux pouvant être hébergés. Ainsi, l'effectif du groupe social et la structure démographique (ratio mâles adultes : femelles adultes : mâles immatures : femelles immatures) peuvent être différents par rapport aux groupes sauvages. De même, certains contextes rencontrés par les animaux en milieu naturel (*i.e.* les déplacements, la recherche alimentaire dans un environnement visuellement obstrué, les phénomènes de fission-fusion ou encore les chasses collaboratives chez les chimpanzés) pouvant être propices aux interactions vocales, sont absents en captivité. A l'inverse, la captivité conduit à l'apparition de contextes nouveaux tels que l'attente

alimentaire précédant les périodes de nourrissage, pouvant aboutir à des phases d'anticipation alimentaire (Waitt et Buchanan-Smith 2001). Toutefois, il est important de préciser qu'à notre connaissance, la captivité peut impacter le comportement vocal (*e.g.* innovation vocale : Hopkins *et al.* 2007 ; apprentissage vocal Wich *et al.* 2009a ; contexte d'utilisation d'une vocalisation : Schaeffer 2011 ; ou référentialité des vocalisations d'alarme : Ouattara *et al.* 2009b), mais n'influence que dans une moindre mesure la socialité (*e.g.* Levé *et al.* 2016).

C'est pourquoi le fait d'aborder la question des échanges vocaux organisés sur le tour de parole chez les grands singes en étudiant à la fois des groupes captifs et des groupes provenant de milieux naturels est essentiel. Ces facteurs socio-contextuels sont en effet susceptibles d'influencer la fréquence, la diversité, la distribution des échanges vocaux ainsi que les règles temporelles et sociales sous-jacentes.

Dans ce travail de thèse, les chimpanzés, gorilles des plaines de l'Ouest et orangs-outans de Bornéo ont été étudiés en captivité en couplant l'approche observationnelle à l'approche expérimentale. Des données provenant de milieux naturels sont venues compléter ce jeu de données pour les chimpanzés et les bonobos. La majorité des étapes de collecte et d'analyse des données présentées par la suite ont été réalisées par mes soins. Le cas échéant, les noms des collaborateurs et leurs tâches ont été précisés.

## 2.1. En parcs zoologiques

Les groupes de chimpanzés, de gorilles et d'orangs-outans captifs ont été étudiés au ZooParc de Beauval (Saint Aignan, France) entre avril et décembre 2018. Des observations se sont également déroulées sur le groupe d'orangs-outans hébergé au Ouwehand Dierenpark (Rhenen, Hollande) entre novembre et décembre 2019.

Les groupes étudiés en captivité ont été constitués de 15 chimpanzés (deux mâles adultes, sept femelles adultes, deux mâles immatures, deux femelles immatures), 11 gorilles des plaines de l'Ouest (un mâle adulte, quatre femelles adultes, trois mâles immatures, trois femelles immatures), 16 orangs-outans de Bornéo (deux mâles adultes, six femelles adultes, cinq mâles immatures, trois femelles immatures).

La structure sociale des groupes dans lesquels évoluaient les chimpanzés et les gorilles des plaines de l'Ouest était similaire à celle décrite en milieu sauvage (voir 1.1.2 et 1.3.2.). En ce qui concerne les orangs-outans de Bornéo, deux stratégies différentes ont été adoptées par les parcs zoologiques pour répondre au mieux aux besoins des groupes hébergés (*e.g.* réduire les agressions et permettre les interactions sociales quand elles semblaient nécessaires à certains

individus). Ainsi, les individus du ZooParc de Beauval étaient regroupés dans le même enclos alors que celui du Ouwehand Dierenpark étaient subdivisés en quatre sous-groupes avec un accès visuel et auditif entre eux. Le tableau 1 détaille les caractéristiques de chaque sujet (sexe, âge, liens de parentés, lieux et conditions d'hébergement).

**Tableau 1 |** Caractéristiques des sujets captifs.

Site : ZPB = ZooParc de Beauval, OD = Ouwehand Dierenpark ;

Ratio âge-sexe = mâles adultes : femelles adultes : mâles immatures : femelles immatures

\* : Individu décédé, - : Parents inconnus, NA : incertitude sur l'identité de l'individu

Site	Espèce	Enclos	Ratio âge-sexe ♂A : ♀A : ♂I : ♀I	Sujet (sous- groupe)	Sexe	Date de naissance – Date d'arrivée (si différentes)	Parents ♀ x ♂		
ZPB	<i>Pan troglodytes</i> spp.	<u>Extérieur :</u> ~1200m <sup>2</sup>	2018 2:7:2:2	Joseph	♂ <sub>3</sub>	1974 ± 1 an	-		
				Charlotte	♀ <sub>0</sub>	01/08/1976	-		
		<u>Intérieur :</u> ~150m <sup>2</sup> x5.30m	Micheline	♀ <sub>0</sub>	1977 ± 1 an	-			
			Baraka	♀ <sub>0</sub>	02/01/1979	-			
			Julie	♀ <sub>0</sub>	1981 ± 1 an	-			
			Bonobo	♀ <sub>0</sub>	1981 ± 1 an	-			
			Gypso	♀ <sub>0</sub>	25/02/1987	-			
			Gamin	♂ <sub>3</sub>	1988 ± 1 an	-			
			Domi	♂ <sub>3</sub>	04/12/1989	-			
			Sangha	♂ <sub>3</sub>	21/07/2006	NA x Bonobo			
			Wamba	♂ <sub>3</sub>	23/11/2008	NA x Bonobo			
			Tumba	♂ <sub>3</sub>	14/01/2009	NA x Domi			
			Lukombe	♂ <sub>3</sub>	03/08/2011	NA x Bonobo			
			Yumbi	♂ <sub>3</sub>	05/02/2014	NA x Bonobo			
			Lobaï	♂ <sub>3</sub>	11/06/2014	NA x Domi			
			<i>Gorilla gorilla gorilla</i>	<u>Extérieur :</u> ~3500m <sup>2</sup>	2018 1:4:3:3	Inge	♀ <sub>0</sub>	02/03/1980	-
						Kabinda	♀ <sub>0</sub>	10/12/1982	-
		<u>Intérieur :</u> ~720m <sup>2</sup> x6m		Sheila	♀ <sub>0</sub>	18/02/1991	-		
				Asato	♀ <sub>0</sub>	20/10/1991	-		
Mayombé	♂ <sub>3</sub>			21/10/2007	Asato x Inge				
Mapenzi	♂ <sub>3</sub>			14/04/2010	Asato x Kabinda				
Kuimba	♂ <sub>3</sub>			19/08/2010	Asato x Tamarilla*				
Sawa	♂ <sub>3</sub>			24/11/2011	Asato x Sheila				
Mayélé	♂ <sub>3</sub>			21/03/2013	Asato Kabinda				
Yamba	♂ <sub>3</sub>			29/12/2015	Asato x Kabinda				
Kiwano	♂ <sub>3</sub>			11/01/2016	Asato x Inge				
<i>Pongo pygmaeus</i> spp.	<u>Extérieur :</u> ~1000m <sup>2</sup>	2018 1:3:3:1	Muda	♂ <sub>3</sub>	29/05/1983 – 20/07/1994	-			
			Christina	♀ <sub>0</sub>	1985/86 – 26/11/1993	-			
	<u>Intérieur :</u> ~165m <sup>2</sup> x5.30m	Manis	♂ <sub>3</sub>	24/09/1990 – 13/04/2010	-				
		Janah	♂ <sub>3</sub>	15/02/1991 – 18/06/1992	-				
		Dayak	♂ <sub>3</sub>	04/03/2009	Muda x Manis				
		Belayan	♂ <sub>3</sub>	22/09/2009	Muda x Janah				
		Sabah	♂ <sub>3</sub>	06/08/2015	Muda x Manis				
Kayan	♂ <sub>3</sub>	06/05/2017	Muda x Janah						
OD	<i>Pongo pygmaeus</i> spp.	<u>Extérieur :</u> ~750m <sup>2</sup>	2019 1:3:2:2	Anak (n°2)	♀ <sub>0</sub>	09/03/1976	-		
				Tjintah (n°3)	♀ <sub>0</sub>	01/05/1984 – 19/08/2005	-		
		<u>Intérieur (pour chaque compartiment) :</u> ~110m <sup>2</sup> x7m	Jewel (n°4)	♀ <sub>0</sub>	1987 ± 1 an – 13/04/1992	-			
			Bako (n°1)	♀ <sub>0</sub>	21/03/1996	-			
			Jingga (n° 2)	♀ <sub>0</sub>	22/10/2007	Bako x Jewel			
			Daya (n°4)	♀ <sub>0</sub>	27/05/2012	Bako x Jewel			
			Sabar (n°4)	♀ <sub>0</sub>	11/05/2016	Bako x Jewel			
			Minggu (n°2)	♀ <sub>0</sub>	10/03/2019	Bako x Tjintah			

Tous les groupes captifs avaient accès à un enclos extérieur si les conditions climatiques le permettaient ainsi qu'à un ou plusieurs enclos intérieurs à température et hygrométrie stables tout au long de l'année. Les enclos intérieurs étaient également connectés à des loges permettant les interactions avec les soigneurs-animaliers. Des modules en bois, de la végétation naturelle

et des bassins agrémentaient ces enclos. Une distribution régulière de nourriture avait lieu six fois par jour. Celle-ci se composait de fruits, de légumes, de protéines animales (*i.e.* œuf et blanc de poulets) ainsi que des aliments composés industriels sous forme de granulés. L'eau était disponible *ad libitum*.

## 2.2. A l'état sauvage

Une collaboration avec l'Université de Zurich nous a permis d'accéder aux données collectées par Maël Leroux (doctorant sous la direction de Dr. Simon W. Townsend) sur la communauté de chimpanzés du site de Budongo (*Pan troglodytes schweinfurthii*) entre février et mai 2019. Au sein de cette communauté d'une soixantaine individus, 35 d'entre eux ont été inclus dans ce manuscrit (12 mâles adultes, 13 femelles adultes, six mâles immatures et quatre femelles immatures, Tableau 2). Un stagiaire de licence 2, Julien Paulet, a pris part au codage des données récoltées sous mon encadrement.

De même, nous avons obtenu des données provenant de la population de bonobos sauvages de Nkala *Forest*. Ces données provenaient d'une campagne de terrain menée de janvier à avril 2012 puis d'octobre à décembre 2012 par Dr. Florence Levréro, Dr. Fabrice Dentressangle, Laure Déruti (tous trois membres du laboratoire ENES, Université Lyon/St Etienne) et Pablo Bombani (de l'ONG Mbou-Mon-Tour). Un total de 19 individus identifiés a été étudiés (six mâles adultes, six femelles adultes, deux mâles immatures, deux femelles immatures et trois individus non-sexés, Tableau 2). FL a analysé les données récoltées sur cette période.

**Tableau 2 |** Caractéristiques des sujets sauvages

Ratio âge-sexe = mâles adultes : femelles adultes : mâles immatures : femelles immatures

\* Individu décédé, ° Incertitude de plusieurs années sur la date de naissance de l'individu, - : Parents inconnus, NA : incertitude sur l'identité de l'individu

Site	Espèce	Taille du territoire	Ratio âge-sexe ♂A : ♀A : ♂I : ♀I	Sujet	Sexe	Date de naissance	Nom de la mère
Budongo- Sonso	<i>Pan troglodytes schweinfurthii</i>	15km <sup>2</sup>	2018 12:13:6:4	NB	♂	1962°	-
				OK	♂	1996 ± 1an	-
				FL	♂	1979°	-
				KL	♂	1979 ± 1an	-
				KU	♂	1979°	-
				KW	♂	1981°	-
				ZF	♀	1982 ± 6mois	BN
				JN	♂	1984 ± 1an	-
				TJ	♂	1984°	-
				SQ	♂	1991 ± 1an	BC
				MS	♀	1992 ± 6mois	NB
				HW	♀	1993 ± 1an	HT
				KT	♂	1993	KU
				SM	♀	1993 ± 1an	SE
				KZ	♀	1995	KW
				ZL	♀	1995	ZN
				PS	♀	1998 ± 1an	PL
				RS	♂	1997	RH
				FK	♀	1999 ± 1an	FL
				RM	♂	2002	RH
				ZD	♂	2003	ZN
				DL	♂	2002 ± 1an	-
				CD	♂	2003 ± 1an	-
				KS	♂	2003	KU
				EV	♂	2004 ± 1an	-
				JS	♂	2006	JN
				KC	♂	2006	KL
				KX	♂	2007	KY
				MB	♂	2009	ML
				TW	♂	2010	TJ
KF	♂	2014	KU				
KO	♂	2014	KL				
IS	♂	2017	IR				
RN	♂	2018	RS				
DB	♂	2018	DL				
Nkala Forest	<i>Pan paniscus</i>	~40km <sup>2</sup>	2012 6:6:2:2	Mpaka	♂	-	-
				Indo	♂	-	-
				Kitoko	♂	-	-
				M6	♂	-	-
				VieuxM	♂	-	-
				Mssmaindroite	♂	-	-
				Mwasi	♂	-	-
				Chinoise	♂	-	-
				Api	♂	-	-
				AFB1	♂	-	-
				AFB2	♂	-	-
				Fssmaindroite	♂	-	-
				Mona	♂	-	-
				JF2	♂	-	-
				JM1	♂	-	Api
				JM2	♂	-	APFB2
				IM1	NA	-	-
				IM2	NA	-	-
IM3	NA	-	-				

### 3. Collecte de données

Au cours de ces recherches doctorales, deux approches distinctes et complémentaires ont été adoptées : l'approche observationnelle et l'approche expérimentale. L'approche observationnelle a eu pour intérêt de décrire les comportements vocaux et sociaux de chaque espèce. Par cette approche, nous avons pu avoir accès aux occurrences de productions vocales,

aux types de vocalisations, à l'organisation temporelle des interactions vocales et leur contexte d'émission, ainsi qu'aux interactions sociales, affiliatives et agonistiques au sein des groupes.

L'approche expérimentale, quant à elle, nous a permis d'appréhender l'importance des règles « conversationnelles » temporelles et sociales pour les individus. Pour cela, nous avons modifié des paramètres clés (*i.e.* sociaux ou temporaux) et conservé les autres variables afin d'observer les effets induits par ces changements.

Le tableau 5, situé en fin de ce chapitre, résume les différentes approches utilisées et les espèces sur lesquelles elles ont été appliquées.

### 3.1. Enregistrements acoustiques et approche observationnelle

L'un des objectifs principaux de ce présent travail était de décrire les interactions vocales des grands singes en vue de mettre en évidence d'éventuels échanges vocaux organisés chez ces derniers. Ce type d'échanges vocaux étant régi par des règles temporelles (*i.e.* respect du tour de parole et évitement actif de la superposition des vocalisations) et sociales (*i.e.* choix non-aléatoire de l'interlocuteur en fonction de ses caractéristiques), nous nous sommes intéressés à ces deux aspects en abordant la production vocale spontanée de ces espèces en contexte intra-spécifique. Cette approche descriptive a été menée sur les populations captives de chimpanzés et d'orangs-outans ainsi que sur les populations sauvages de chimpanzés et de bonobos.

Par la suite, le principe général des méthodes d'enregistrements vocales et des méthodes d'observations comportementales est présenté, puis abordé plus en détails dans les **articles 3 et 4**.

#### 3.1.1. La production vocale

##### *a. Méthodes d'enregistrements et d'extractions*

Les données vocales obtenues en captivité sur les chimpanzés et les orangs-outans ont été collectées en suivant la méthode d'échantillonnage en toutes occurrences (*all occurrence* : Altmann, 1974) par session de 1 heure à 1 heure 30 minutes réparties aléatoirement dans la journée entre 9 heures à 18 heures (temps d'enregistrement total par espèce : chimpanzés = 26 heures ; orangs-outans du ZooParc de Beauval = 17 heures ; orangs-outans du Ouwehand Dierenpark = 25 heures). Les vocalisations ont été enregistrées à une distance maximale de 20 mètres à l'aide d'un enregistreur portatif Marantz® PMD660 (taux d'échantillonnage 44,1 Hz, résolution 32 bits, format .wav) connecté à un microphone Sennheiser® MK70-1 ultra-directionnel.



Les contraintes en milieu naturel ont nécessité des méthodes d'enregistrement différentes. En ce qui concerne les bonobos, du fait que la communauté n'était pas encore pleinement habituée à la présence des scientifiques au moment de la collecte, les données vocales ont été reportées par un échantillonnage *ad libitum* (Altmann 1974) par le biais d'un cahier de terrain ou à l'oral via un enregistreur audio SONY®. Elles représentent 91 heures 04 minutes de contact visuel et/ou auditif avec les primates, l'accès visuel aux bonobos n'étant pas toujours permis par la densité du couvert végétal.

A l'inverse, les chimpanzés de la communauté de Budongo sont étudiés en continu depuis 1990, et donc habitués à la présence des scientifiques. Cela a permis la collecte de données vocales à une distance moyenne de 10 mètres des individus (avec des variations importantes dues au couvert végétal) en suivant la méthode d'échantillonnage de focal par animal (*focal individual sampling* : Altmann, 1974) sur les adultes et en échantillonnage *ad libitum* (Altmann 1974) sur les immatures. Un total de 44 heures et 34 minutes d'enregistrements par session de 2 heures réparties de 9 heure à 18 heure a pu être mené à l'aide d'un enregistreur Marantz® PMD661 mk3 (taux d'échantillonnage 44,1 Hz, résolution 32 bits, format .wav) connecté à un microphone Sennheiser ME66/K6 ultra-directionnel.

Les données vocales provenant des populations captives et de la communauté de chimpanzés sauvages ont toutes été extraites à l'aide du logiciel Audacity® 2.4.2. puis traitées par inspection visuelle à l'aide des spectrogrammes générés par le logiciel PRAAT® 6.1.09. respectant une fenêtre d'analyse gaussienne. De plus, le contexte de production était notifié en fonction de l'état comportemental de l'émetteur au moment de l'émission vocale, et ce aussi bien en captivité qu'en milieu naturel (Tableau 3). Les données obtenues dans la communauté de bonobos de Nkala *Forest* ont été extraites par lecture des documents papiers et audio issus de la campagne de terrain de 2012. Les contextes de production n'étaient pas connus.

**Tableau 3 |** Description des différents contextes rencontrés en milieu captif et sauvage.  
 + : présence de ce contexte dans le milieu donné ; - : contexte non-observé dans le milieu donné

Contexte	Définition	Milieu captif	Milieu sauvage
Affiliatif	L'émetteur est impliqué dans un contact physique social positif avec un autre individu. Les comportements de jeux sont inclus dans le contexte affiliatif.	+	+
Agonistique agressive	L'émetteur chasse ou attaque physiquement un autre individu.	+	+
Agonistique de soumission	L'émetteur évite calmement un individu ou s'enfuit à son approche	+	+
Alimentation	L'émetteur cherche, prépare et ingère de la nourriture (solide ou liquide), cela inclut l'allaitement.	+	+
Attente alimentaire	L'émetteur a vu et/ou entendu les soigneurs animaliers dans les couloirs avec la nourriture avant sa distribution.	+	-
Chasse	L'émetteur poursuit, attrape et met à mort une espèce cible comprend les patrouilles ( <i>i.e.</i> les individus marchent les uns derrière les autres en ligne), mais n'inclut pas la consommation ultérieure de viande.	-	+
Déplacement arboricole vertical	L'émetteur monte ou descend entre le sol et la canopée.	-	+
Déplacement au sol	L'émetteur se déplace d'une zone à l'autre au sol, n'inclut pas les locomotions brèves entre individus au sein du groupe.	+	+
Déplacement arboricole en canopée	L'émetteur se déplace d'une zone à l'autre dans la canopée, n'inclut pas les locomotions brèves entre individus au sein du groupe.	-	+
Frustration	L'émetteur est confronté à un autre individu qui refuse de partager sa nourriture ou d'interagir sexuellement	+	+
Invitation sexuelle	L'émetteur présente ses organes génitaux, inclut les accouplements.	+	+
Quémande alimentaire	L'émetteur demande l'accès à la nourriture à un autre individu	+	+
Rencontre inter-communauté	L'émetteur fait partie d'une patrouille de groupe ( <i>i.e.</i> les individus marchent les uns derrière les autres en ligne) se déplaçant à la rencontre d'une communauté voisine, ou l'émetteur se joint au groupe lors d'une rencontre déjà en cours.	-	+
Rencontre intra-communauté	L'émetteur rejoint ou est rejoint par un individu ou un groupe de sa propre communauté.	+	+
Repos	L'émetteur reste sur place, sans être engagé dans des comportements autodirigés (incluant l'auto-toiletage) ou d'autres activités physiques. L'émetteur peut aussi être en train de dormir.	+	+

### *b. Types de vocalisations*

La classification des vocalisations répertoriées chez les chimpanzés (*i.e.* captifs et sauvages) et les orangs-outans a été réalisée par inspection visuelle des sonogrammes sur la base de répertoires vocaux précédemment établis. Ainsi, le répertoire vocal décrit par Hardus et ses collègues en 2009 chez les orangs-outans a servi de base pour la détermination des vocalisations produites par les individus étudiés (Tableau 4a). Chez les chimpanzés, la classification vocale a été réalisée avec le répertoire vocal décrit par Slocombe et Zuberbühler (2010), couplée à la description proposée par Goodall en 1986 (Tableau 4b). Lorsque différentes vocalisations étaient combinées, le terme de *mixed-call* était employé (Lemasson *et al.* 2018).

Les vocalisations produites par les bonobos sauvages n'ont pas été classées au moment de leur collecte par manque d'informations. En effet, les enregistrements audio des vocalisations n'étaient pas systématiques, et le déplacement fréquent des individus réduisait la probabilité d'obtenir des enregistrements de bonne qualité.

**Tableau 4a** | Répertoire vocal des orangs-outans de Bornéo utilisé au cours de cette étude et les principaux contextes d'émission de ces vocalisations en milieu naturel. La description des différents contextes provient des observations menées par Hardus et ses collègues en 2009.

Type de vocalisation	Principaux contextes de production	Type de vocalisation	Principaux contextes de production
<i>Ahh</i>	Lors de jeux sociaux	<i>Long call</i>	Produit uniquement par les mâles adultes pour : attirer les femelles, avertir/repousser les rivaux – Lors d'une perturbation proche (sociale ou non)
<i>Ahoor call</i>	Lors d'une interaction agonistique	<i>Lork</i>	Lors d'une rencontre avec un potentiel danger (e.g. humain, prédateur ou congénère)
<i>Bared teeth scream</i>	Lors d'une agression par un congénère	<i>Mating squeal</i>	Lors d'une copulation
<i>Bark</i>	Lors d'une interaction agonistique	<i>Nest smack</i>	Lors de la préparation du nid
<i>Chomp</i>	Avant une interaction agonistique	<i>Play ooh</i>	Lors de jeux solitaires
<i>Complexe call</i>	Lors d'une intimidation	<i>Raspberry</i>	Lors de la préparation du nid
<i>Crying/Screaming</i>	Lors d'une séparation mère/jeune – Lors d'une agression par un congénère	<i>Roar</i>	Lors d'une interaction agonistique
<i>Fast long call</i>	Lors d'une perturbation proche (sociale ou non)	<i>Rolling call</i>	Lors d'une intimidation
<i>Frustration scream</i>	Lors d'une quémante alimentaire	<i>Squeak</i>	Lors d'une copulation – Lors d'une situation anxiogène
<i>Gorkum</i>	Lors d'une intimidation	<i>Throatscrape</i>	Lors de la réunion mère/enfant Lors de la préparation du nid En présence d'un mâle adulte
<i>Grumble</i>	Lors d'une rencontre avec un observateur humain – Lors d'une copulation – Lors d'une interaction agonistique	<i>Whimper</i>	Lors d'une situation anxiogène
<i>Grumph</i>	Lors d'une intimidation - Lors d'une rencontre avec un observateur humain		
<i>Grunt</i>	Lors de la réunion mère/enfant Lors de la préparation du nid En présence d'un mâle adulte		
<i>Kiss squeak</i>	Lors d'une rencontre avec un potentiel danger (e.g. humain, prédateur ou congénère)		

**Tableau 4b** | Répertoire vocal des chimpanzés de la communauté de Budongo utilisé au cours de cette étude et les principaux contextes d'émission de ces vocalisations en milieu naturel. La description des différents contextes provient des observations menées par Slocombe et Zuberbühler en 2010 <sup>(a)</sup> et complété par celui de Goodall en 1986 <sup>(b)</sup>.

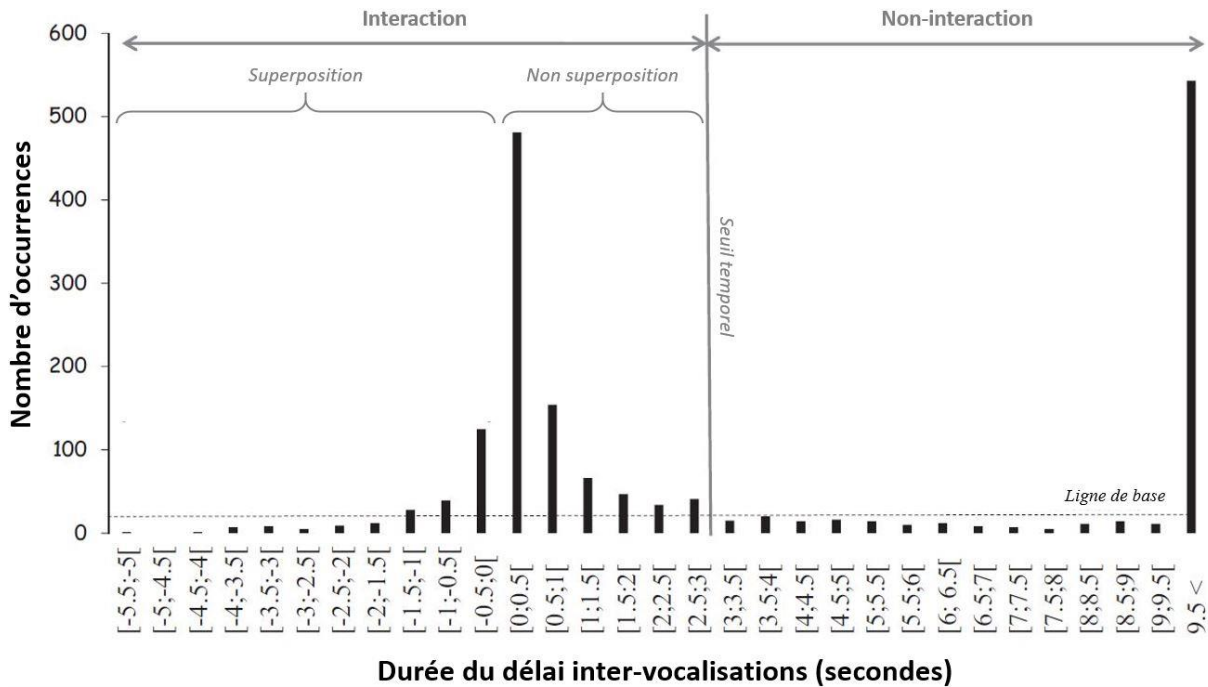
Type de vocalisation	Principaux contextes de production	Type de vocalisation	Principaux contextes de production
<i>Bark</i> <sup>ab</sup>	Lors d'une interaction agonistique – Lors d'une phase d'excitation sociale	<i>Pant-hoot</i> <sup>ab</sup>	Lors d'une intimidation – Lors d'un déplacement – Lors de l'approche d'un site d'alimentation – Lors d'une phase d'excitation sociale – Lors d'une rencontre avec un autre groupe d'individus
<i>Cough</i> <sup>ab</sup>	Lors d'une intimidation	<i>Rough grunt</i> <sup>ab</sup>	Lors de l'approche d'un site d'alimentation – Avant de s'alimenter
<i>Grunt</i> <sup>a</sup>	Lors des phases de repos – Lors de la préparation du nid	<i>Scream</i> <sup>ab</sup>	Lors d'une interaction agonistique – Lors d'une phase d'excitation sociale – En présence d'un danger extrême (e.g. attaque sévère ou prédateur)
<i>Huu</i> <sup>ab</sup>	Lors d'une perturbation proche (sociale ou non) – Lors d'une rencontre avec un potentiel danger (e.g. humain, prédateur)	<i>Squeak</i> <sup>ab</sup>	Suite à une intimidation – Lors d'une copulation
<i>Laughter</i> <sup>ab</sup>	Lors de jeux sociaux	<i>Waa-bark</i> <sup>ab</sup>	Lors d'une intimidation
<i>Pant</i> <sup>ab</sup>	Lors d'une rencontre sociale avec un supérieur hiérarchique – Lors d'une phase de toilettage social	<i>Whimper</i> <sup>ab</sup>	Lors d'une séparation mère/jeune – Lors d'une quémante alimentaire
<i>Pant-grunt</i> <sup>ab</sup>	Lors d'une rencontre sociale avec un supérieur hiérarchique	<i>Hoo</i> <sup>b</sup>	Lors d'une phase de repos proche des congénères

### c. Organisation temporelle des productions vocales

#### Cas des échanges vocaux organisés sur le tour de parole

Pour caractériser les interactions vocales, et tout particulièrement les éventuels échanges vocaux organisés des populations étudiées, nous nous sommes intéressés à la dimension temporelle de ce type de production vocale. En effet, l'une des principales caractéristiques temporelles de ce type d'interaction dit « conversationnel » est l'évitement actif de la superposition des vocalisations qui se manifeste par le respect d'un délai inter-vocalisations court. Ce délai inter-vocalisations se traduit par un intervalle de silence court séparant deux vocalisations consécutives, émises par deux individus distincts. Ce délai est comptabilisé à partir de la fin de la première vocalisation jusqu'au début de la seconde. Le type de vocalisation n'a pas d'importance à ce stade. Néanmoins, un seuil temporel maximal à ce délai inter-vocalisations devait être déterminé pour définir les vocalisations produites au sein d'un même échange vocal. Pour cela, nous avons choisi d'opter pour une méthode d'analyse graphique (e.g. Figure 5) similaire à celle utilisée pour mettre en évidence des échanges vocaux

organisés chez un groupe captif de gorilles des plaines de l’Ouest (Lemasson *et al.* 2018) et d’un groupe captif de bonobos (Levréro *et al.* 2019), et ayant déjà fait ses preuves chez des primates non-Hominidés (*e.g.* mone de Campbell Lemasson *et al.* 2010).



**Figure 5** | Exemple de représentation graphique de la distribution de la durée des délai inter-vocalisations de différents émetteurs inspiré de l’article de Lemasson et collaborateurs (2018). Le seuil indiquant le délai de réponse maximale (ligne verticale grise) pour que les vocalisations soient considérées comme faisant partie d’une interaction vocale a été positionné à l’endroit où les délais inter-vocalisations commencent à s’élever au-dessus de la ligne de base (ligne pointillée grise) qui représente le « bruit de fond » des délais inter-vocalisations hors interactions vocales.

Dans le cas où plus de deux vocalisations se chevauchent (*i.e.* durée du délai inter-vocalisations négatives), il a été conservé les vocalisations pour lesquelles les émetteurs étaient identifiés. Une des limites de cette méthodologie réside dans l’impossibilité de distinguer les chevauchements synchrones, pouvant faire référence à des *chorus* sociaux, par rapport aux simples superpositions ne pouvant être considérées comme de « vraies » interactions vocales (Ravignani *et al.* 2014).

**Productions vocales à l’échelle de l’individu**

Les productions vocales d’un individu peuvent suivre deux organisations temporelles, qu’elles soient ou non produites au cours d’une interaction vocale : une vocalisation peut être produite seule ou en série (*i.e.* de façon consécutive et séparée par de courts intervalles

silencieux). Le même seuil que celui utilisé pour définir un échange vocal a été utilisé pour définir si une vocalisation est produite seule ou en série (Koda 2004; Lemasson *et al.* 2013a, 2018).

### 3.1.2. Facteurs socio-démographiques et contextuels

Pour comprendre les bases sociales des échanges vocaux, il est crucial de chercher à appréhender les relations sociales existantes entre les individus. Dans un souci de cohérence avec les précédentes études menées sur les échanges vocaux chez les grands singes (*i.e.* Lemasson *et al.* 2018; Levréro *et al.* 2019) et pour permettre la comparaison des résultats obtenus, nous avons adopté des méthodes similaires de collecte et d'analyse des réseaux sociaux. Cette collecte a été menée sur les groupes captifs de chimpanzés (**article 3**) et d'orangs-outans (données non-présentées dans ce manuscrit car les caractéristiques de leurs productions vocales ne nous ont pas permis de les exploiter) à l'aide d'un micro-cravate AV-JEFE® TCM 160 connecté à un enregistreur portatif Marantz® PMD660 et d'un carnet de notes.

Pour cela, la méthode d'échantillonnage de *focal* par animal (Altmann 1974) a été adoptée afin de collecter l'ensemble des comportements sociaux affiliatifs, agonistiques agressifs et de soumission en utilisant le répertoire comportemental mis au point par Pollick et De Waal (2007). Cette méthode est fréquemment utilisée pour étudier les relations sociales chez les primates non-humains (*e.g.* capucin à face blanche *Cebus capucinus* : Perry 1996; babouin *Papio hamadryas ursinus* : Silk *et al.* 2003; mone de Campbell : Lemasson *et al.* 2006; gorille des plaines de l'Ouest : Lemasson *et al.* 2018; macaque : McFarland et Majolo 2011; chimpanzé : Funkhouser *et al.* 2018). Cet échantillonnage était mené par session de 10 minutes par individu répartie équitablement de 9 heure à 18 heure pour un total de 1 heure et 30 minutes pour chaque animal. En complément de cette méthode d'échantillonnage, les distances inter-individuelles séparant les membres du groupe ont été reportées par le biais de la méthode d'échantillonnage instantanée sur l'ensemble du groupe (*i.e.* *scan sampling* : Altmann 1974). L'échantillonnage était réalisé toutes les 10 minutes, pendant des sessions de 1 heure et 30 minutes réparties de 9 heure à 18 heure.

Les données sociales obtenues grâce aux deux méthodes citées précédemment ont permis de calculer trois indices sociaux individuels que sont : (i) l'indice de dominance (Lemasson et Hausberger 2004; Zumpe et Michael 1986), (ii) le niveau d'intégration sociale par le biais de l'indice d'association dyadique proposé par Prieur et collègues (2016a), et (iii) la

proximité spatiale (Lemasson *et al.* 2018). Deux autres données démographiques étaient également systématiquement relevées : l'âge et le sexe des individus.

### 3.2. Approche expérimentale

Les expériences de repasse de vocalisations sont un moyen intéressant de questionner la fonctionnalité et/ou la pertinence d'un stimulus pour l'individu en enregistrant les réponses comportementales induites par la diffusion de la stimulation dans un contexte donné (Zuberbühler et Wittig 2011). L'approche expérimentale permet, entre autres, de réduire l'ambiguïté sur la relation causale entre les stimuli et les réponses comportementales, d'étudier systématiquement des événements rares et de tester des hypothèses en contrôlant systématiquement les variables (Zuberbühler et Wittig 2011).

Au cours de ce projet de recherche, deux séries d'expérimentations aux finalités distinctes devaient être réalisées sur le groupe de gorilles des plaines de l'Ouest du ZooParc de Beauval. Une première a été menée entre avril et décembre 2018 et une seconde de janvier à juin puis de septembre à octobre 2020. Cette seconde expérimentation a fait l'objet du stage de master 2 de Thomas Ferrari, dont j'ai été le co-encadrant, mais n'a pas pu aboutir en raison de la crise sanitaire. Dans ce manuscrit, seule la première série d'expérimentation est présentée (**article 2**).

Aucune expérience n'a été menée chez les chimpanzés et les orangs-outans car les caractéristiques de leurs productions vocales ne nous ont pas permis de développer cette approche.

#### 3.2.1. Paradigme de violation des attentes

Le principe du paradigme de violation des attentes est de comparer la réaction spontanée d'un (ou plusieurs) individu(s) suite à la présentation d'une situation habituelle en opposition à une situation inattendue (Zuberbühler et Wittig 2011). La réponse comportementale spontanée à cette violation peut ainsi nous informer quant à la capacité des sujets à percevoir et discriminer ces deux situations sur la base du paramètre que l'expérimentateur a préalablement choisi de modifier. La réponse comportementale se traduit majoritairement par une modification de l'état attentionnel de l'individu testé (*e.g.* Bergman *et al.* 2003; Bouchet *et al.* 2017; Lemasson *et al.* 2011; Reber *et al.* 2019). Cette modification de l'état attentionnel peut se manifester par une modification de l'orientation du regard, de la tête, du corps, ou encore par un arrêt des activités en cours.

Avec ce type de paradigme expérimental, la direction de la réponse comportementale des sujets testés peut aussi bien dépendre de la stimulation que du contexte. Traditionnellement, une réponse comportementale plus forte de la part du sujet envers la situation inattendue que pour une situation habituelle témoigne de la surprise de l'individu (Lewis et Goldberg, 1969). Néanmoins, de nombreuses d'études menées sur les primates humains et non-humains révèlent une réponse comportementale plus importante envers une situation cohérente ou familière que face à l'incongruence d'une condition (*e.g.* Ghazanfar et Logothetis 2002; Kuhl et Meltzoff 1982; Lemasson *et al.* 2011; Meltzoff et Borton 1979; Reber *et al.* 2019).

### 3.2.2. Objectif de l'expérimentation

L'étude observationnelle menée par Lemasson et collaborateurs (2018) révèle l'existence d'échanges vocaux au sein d'un groupe captif de gorilles des plaines de l'Ouest hébergés au ZooParc de Beauval. Au cours de ces observations, les interactions vocales apparaissent préférentiellement entre individus d'âges proches (*i.e.* possible règle sociale) tout en respectant un délai inter-vocalisations court (*i.e.* possible règle temporelle). En nous basant sur ces premiers résultats observationnels, notre objectif était de tester la pertinence de ces éventuelles règles sociales (*i.e.* interaction entre individus d'âges proches) et temporelles (*i.e.* respect du délai inter-vocalisations) pour l'audience, et ce par le biais d'un paradigme de violation des attentes.

Pour ce faire, nous avons comparé la réaction spontanée de six individus, provenant du groupe étudié par Lemasson et collaborateurs en 2018, à l'encontre d'échanges vocaux simulés de leurs congénères qui respectaient ou non le délai inter-vocalisations court et la proximité d'âge des interlocuteurs. Ces échanges vocaux simulés ont été constitués, par mes soins, à partir de vocalisations que j'ai préalablement enregistrées puis traitées à l'aide des logiciels Audacity® 2.4.2, PRAAT® 6.1.092 et GoldWave® 6.52. Pour appréhender l'intérêt des sujets testés vis-à-vis des stimulations présentées, les comportements d'attention visuelle, tels que l'orientation de la tête, les déplacements et la durée du premier regard vers le stimulus ont été relevés.

### 3.2.3. Dispositif expérimental

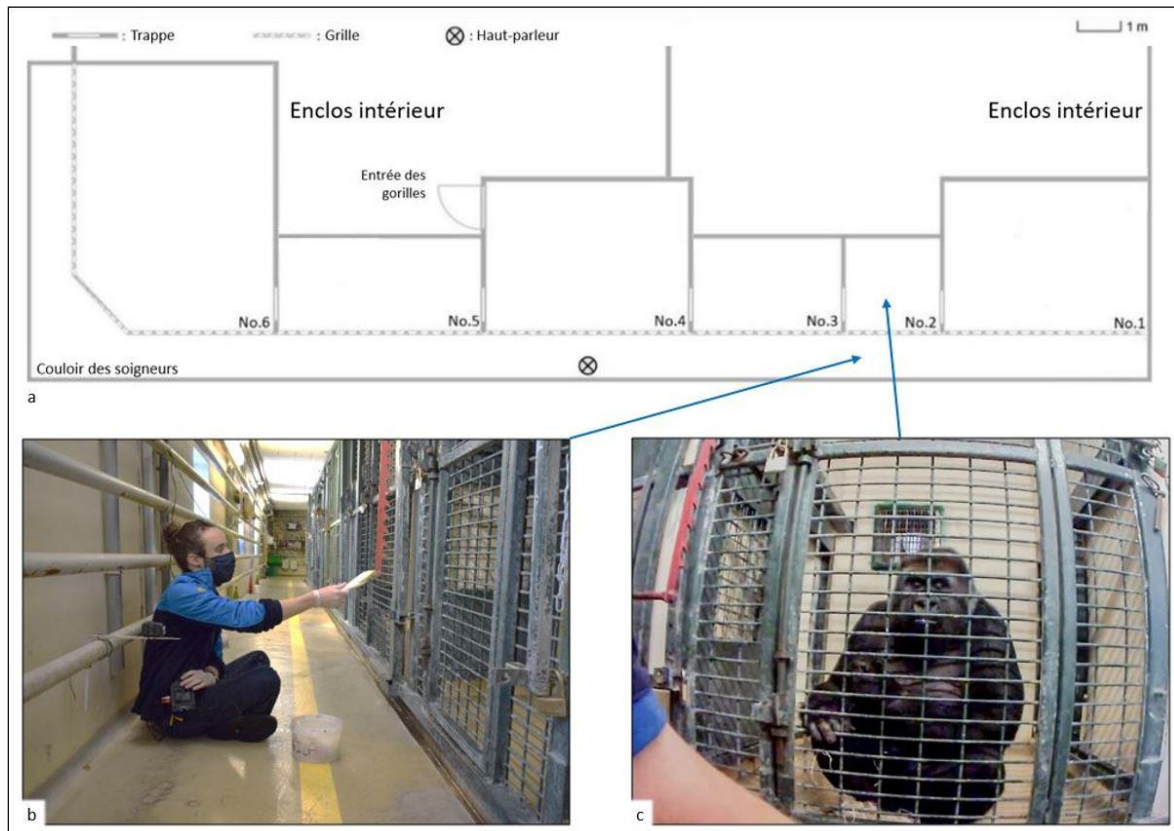
L'expérimentation s'est déroulée dans les loges des animaux entre 8 heure et 10 heure, au moment du nourrissage. Cette période de la journée constituait une période privilégiée durant laquelle les soigneurs et l'expérimentateur étaient en mesure d'interagir directement avec les gorilles à travers une grille. Les individus se répartissaient de manière routinière et volontaire dans six loges en fonction de leurs liens de parentés (*i.e.* la mère avec ses jeunes). Une fois les



individus positionnés, les soigneurs étaient en mesure de refermer les trappes séparant les différentes loges, laissant un accès visuel et auditif entre les loges (Figure 6).

Pour chaque test, le stimulus auditif était diffusé à l'aide d'un haut-parleur TAG® Premio 8.2 fixé à une intensité 60 dB SPL (*i.e.* intensité correspondante à celle des vocalisations de contact enregistrées dans ce groupe dans des conditions similaires et à une distance moyenne de 2 mètres), invisible au sujet testé et connectée à un enregistreur portatif Marantz® PMD660 (taux d'échantillonnage 44.1kHz, résolution 32bits, format .wav). Le haut-parleur était positionné dans le couloir avant l'entrée des animaux. La position du haut-parleur devait être pertinente par rapport à la stimulation diffusée (*i.e.* face à la loge du mâle adulte, si une vocalisation de ce dernier devait être diffusée).

Tout au long de l'expérimentation, l'expérimentateur a adopté une position standardisée pour le nourrissage des animaux. La scène était filmée par une caméra positionnée derrière l'expérimentateur (Figure 6).



**Figure 6** | Mise en place du dispositif expérimental. Représentation schématique des six loges de nuit dans lesquelles avaient lieu l'ensemble des tests (a), l'expérimentateur prenait place face à l'animal et adoptait une position standardisée en ayant la possibilité de déclencher la stimulation dans la fenêtre temporelle adéquate (b), les comportements du sujet étaient filmés à l'aide de la caméra positionnée proche de l'expérimentateur (c).

### 3.3. Considérations statistiques

La normalité et l’homogénéité des variances étaient systématiquement évaluées en inspectant les résidus par le biais du test de Shapiro-Wilk. La distribution des données ne suivant pas la loi normale, nous avons effectué uniquement des tests non-paramétriques avec un seuil d’erreur  $\alpha$  fixé à 95% (Siegel et Castellan, 1998). Les données ne pouvant être analysées par le biais de tests statistiques ont fait l’objet d’une description. L’ensemble des données a été traité à l’aide du logiciel R Studio 0.99.903 R version 3.3.1: R Core Team2016).

**Tableau 5** | Récapitulatif des différentes approches menées au cours de ce projet de recherche doctorale.

Il est précisé l’article dans lequel chaque approche a été utilisée

- : approche non-utilisée, \* Données non-présentées dans ce manuscrit, <sup>○</sup> Données collectées par un tiers

	Populations captives			Populations sauvages		
	Orangs-outans de Bornéo	Gorilles des plaines de l’Ouest	Chimpanzés	Chimpanzés	Bonobos	
Approche observationnelle	Données sociales	*	-	Article 3	-	-
	Données vocales	Article 4	-	Article 3	Article 4 <sup>○</sup>	Article 4 <sup>○</sup>
Approche expérimentale	-	Article 2	-	-	-	-



# Chapitre III

## **Pertinence sociale des règles d'interactions vocales**





# Article 2

## **Breaking conversational rules matters to captive gorillas: A playback experiment**

Loïc Pougnault<sup>1,2,3</sup>, Florence Levréro<sup>2</sup>, Baptiste Mulot<sup>3</sup>, Alban Lemasson<sup>1</sup>

<sup>1</sup> Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000 Rennes, France.

<sup>2</sup> Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES / CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France.

<sup>3</sup> ZooParc de Beauval & Beauval Nature, 41110 Saint Aignan, France

*Article publié dans Scientific Reports*

<https://doi.org/10.1038/s41598-020-63923-7>



## Synthèse : article 2



Chez les gorilles des plaines de l'Ouest, les règles “conversationnelles” sont-elles socialement pertinentes pour l'audience ?

### Contexte

Une étude récente en captivité a mis en lumière l'existence d'échanges vocaux fondés sur le principe de tour de parole avec des règles « conversationnelles » temporelles et sociales chez les gorilles des plaines de l'Ouest. Ces échanges vocaux rassemblent majoritairement des interlocuteurs d'âge proche qui, de plus, respectent un délai inter-vocalisations court prévenant ainsi toutes coupures de parole. Au cours de notre propre recherche, nous avons tenté de démontrer la pertinence pour ces gorilles de ces dites « règles » par le biais d'un paradigme de violation des attentes.

### Méthode

#### ① Constitution des stimuli

Pré-enregistrements : 125 échanges vocaux spontanés



Création de trois conditions expérimentales

Echange vocal habituel :

- Respect du délai inter-vocalisations
- Respect de la proximité d'âge



Echange vocal brisant la règle temporelle :

- Violation du délai inter-vocalisations
- Respect de la proximité d'âge



Echange vocal brisant la règle sociale :

- Respect du délai inter-vocalisations
- Violation de la proximité d'âge



#### ② Diffusion des stimuli

N<sub>testé</sub> = 3 adultes & 3 immatures



Chacun soumis aux trois conditions dans un ordre aléatoire

#### ③ Analyses des comportements attentionnels



1 min avant diffusion | 1 min après diffusion

Diffusion

Focal par individu

- Durée totale d'orientation de la tête vers le haut-parleur
- Durée de la première orientation de tête vers le haut-parleur
- Nombre de déplacements vers le haut-parleur

### Principaux résultats

- Un degré d'attention plus important pour un échange vocal habituel que pour un échange vocal violant la règle temporelle.
- Un degré d'attention équivalent porté aux échanges vocaux habituels et à ceux violant la règle sociale.
- Une influence de l'âge de l'individu testé sur l'intérêt porté à ces deux patterns de violation.

### Conclusion

Le faible intérêt que portent les sujets testés à l'échange vocal brisant la règle temporelle, en comparaison à un échange la respectant, témoigne de l'importance qu'ils accordent au respect du délai inter-vocalisations. La pertinence de la différence d'âge entre interlocuteurs pour les gorilles est nettement moins claire. Pourtant, l'attention accrue des jeunes sujets envers un échange impliquant un interlocuteur de leur classe d'âge suggère que, même si ce n'est pas une règle absolue, les gorilles peuvent avoir des attentes à cet égard. Par ailleurs, l'intensité de l'attention des sujets vis-à-vis des stimuli habituels varie également en fonction de leur âge, ce qui peut sous-entendre un éventuel rôle de l'apprentissage social dans la perception de ces règles d'interactions vocales.





OPEN

# Breaking conversational rules matters to captive gorillas: A playback experiment

Loïc Pougault<sup>1,2,3</sup>✉, Florence Levréro<sup>2,4</sup>, Baptiste Mulo<sup>3</sup> & Alban Lemasson<sup>1,4</sup>

Across human cultures, conversations are regulated by temporal and social rules. The universality of conversational rules suggests possible biological bases and encourages comparisons with the communicative interactions of nonhuman animals. Unexpectedly, few studies have focused on other great apes despite evidence of proto-conversational rules in monkeys, thus preventing researchers from drawing conclusions on potential evolutionary origins of this behaviour. A previous study showed however that western lowland gorillas engage in soft call interactions that seem temporally- and socially-ruled. Indeed, interactions occurred mainly between individuals close in age who followed a preset response delay, thus preventing call overlap. Here, we experimentally investigated the presence of these rules in a captive gorilla group, using a violation-of-expectation paradigm. Head orientation responses suggest that the respect of response delay matters to subjects, but the importance of the interlocutors' age proximity appeared less clear. The intensity of the response varied with subjects' age in a context-dependent way, supporting a possible role of learning. Our findings support the growing number of studies highlighting the importance of vocal turn-taking in animals and a possible sociogenesis of this ability. The capacity to “converse” might have been a key step in the co-evolution of communication and complex sociality.

Despite the diversity of human cultures, some basic conversational rules are respected in all societies and thus appear universal<sup>1</sup>. These features are gathered in a so-called “contract of communication” that conversing interlocutors informally agree on<sup>2</sup>. The contract takes into account “context relevance” (*i.e.* the evaluation of the context as pertinent or not to initiate a conversation), “reciprocity” (*i.e.* the evaluation of both partners as valid interlocutors), and “contract-based temporal rules” (*i.e.* the respect of a reciprocal exchange of alternating, short, and flexible turns between two or more interlocutors<sup>3</sup>, and speech overlap avoidance). Several authors have questioned the possible biological bases of these conversational rules and have, for several decades, conducted cross-species comparisons to understand their origin and role across animal communication systems<sup>4–7</sup>.

To facilitate cross species investigations, authors typically distinguish “conversation-like vocal exchanges” from other vocal patterns like isolated calling (one call emitted and no other calls can be heard around, *e.g.* in red-capped mangabeys *Cercocebus torquatus*<sup>8</sup>), repeated calling (the same caller calls several times in a row, *e.g.* in Japanese macaques *Macaca fuscata*<sup>9</sup>), disorganized phonoresponses (one individual produces a call, typically an alarm call that triggers calls in an apparent chaotic way from the other group members, *e.g.* in blue monkeys *Cercopithecus mitis stuhlmanni*<sup>10</sup>), chorusing (two or more individuals overlap their emission of a given call type, *e.g.* in chimpanzees *Pan troglodytes*<sup>11</sup>) and duetting (two individuals synchronise long series of calls or songs with stereotyped temporal association: *e.g.* in birds<sup>12–14</sup>, and gibbons *Hylobates syndactylus*<sup>15</sup>). “Conversation-like vocal exchanges” are distinguished from the other types of vocal interactions because of their following specific features. First, they involve a diversity of recurrent vocal partners (which differs from duets). Second, interlocutors are typically familiar individuals belonging to a given social group and can be of any age or either sex (which differs from synchronized signalling<sup>16–20</sup>). Third, “conversation-like vocal exchanges” are not restricted to a specific context, which differs from – usually long-distance – collective communication associated with mate attraction, territory protection and environmental disturbance<sup>21</sup>, and time of the day or year (excluding morning choruses and reproductive signals<sup>22,23</sup>). Most commonly, these “conversation-like vocal exchanges” are observed

<sup>1</sup>Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000, Rennes, France.

<sup>2</sup>Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES/CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France. <sup>3</sup>ZooParc de Beauval & Beauval Nature, 41110, Saint Aignan, France. <sup>4</sup>These authors contributed equally: Florence Levréro and Alban Lemasson. ✉e-mail: [loic.pougault@univ-rennes1.fr](mailto:loic.pougault@univ-rennes1.fr)

in everyday and basically all day long short-distance communication (*i.e.* soft calls exchanged in peaceful interactions<sup>5</sup>), as in humans<sup>3,24</sup>.

The “conversation-like vocal exchanges” are temporally organized on the basis of turn-taking which refers to “the orderly exchange of purely communicative signals or behaviours between individuals characterized by principles for the coordination turn transfer, which result in observable temporal regularities” (cited from Pika and colleagues<sup>25</sup>). Communicative turn-taking has been found in a broad range of nonhuman species (*e.g.* elephants *Loxodonta africana*<sup>26</sup>; sperm-whales *Physeter macrocephalus*<sup>27</sup>; common bottlenose dolphins *Tursiops truncatus*<sup>28</sup>; starlings *Sturnus vulgaris*<sup>4</sup>; naked mole-rats *Heterocephalus glaber*<sup>29</sup>; bats *Diaemus youngi*<sup>30</sup>), including a broad range of nonhuman primates (NHP) species<sup>25,31,32</sup>. Thus, to respect turn-taking, interlocutors have to obey basic rules<sup>5</sup>. Interlocutors have coordinated roles, one of them sends a first signal and the other responds, and, when the conversation keeps going, the two interlocutors alternate. The respondent typically adjusts the timing of its vocal response by respecting a minimum silence gap (to prevent call overlap) and a maximum silence gap (to ensure call coordination). A recent study on marmosets, focusing on long periods of vocal interactions, showed such temporal coordination of vocal productions between individuals<sup>32</sup>.

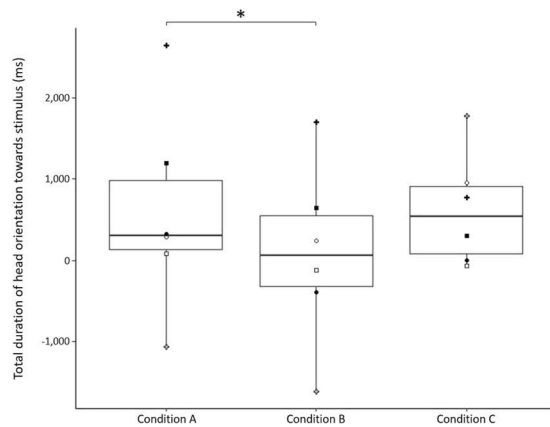
Another important rule shared by human, nonhuman primates and several other animal species is that exchanging partners are not randomly nor opportunistically chosen (bottlenose dolphins<sup>28</sup>, large-billed crows *Corvus macrorhynchos*<sup>33</sup>, meerkats *Suricatta suricatta*<sup>34</sup>, sperm whales<sup>27</sup>, elephants<sup>35</sup>). Dunbar<sup>36</sup> suggested that monkey call exchanges had started to function as a mean to groom-at-distance when the size of group became too large to allocate enough time to physical grooming among all members and maintain close bonds. Indeed, “conversation-like vocal exchanges” occur mainly between two interlocutors that are determined, according to the study species, by their affiliative bonds (humans<sup>37</sup>; NHP: spider monkeys *Ateles geoffroyi*<sup>38</sup>, Japanese macaques<sup>39</sup>, pygmy marmosets *Cebuella pygmaea*<sup>6</sup>, bonobos *Pan paniscus*<sup>40</sup>) or by their respective ages (humans<sup>41</sup>; NHP: common marmosets<sup>42</sup>, western lowland gorillas *Gorilla gorilla gorilla*<sup>43</sup>, Campbell’s monkeys<sup>31</sup>). In some cases, it was even found that human and nonhuman primates advertise this preferential bond by responding to one another using a matching acoustic structure, a phenomenon known as “vocal convergence” in ethology or “vocal accommodation” in sociolinguistics (humans<sup>44</sup>; NHP: spider monkeys<sup>38</sup>, Diana monkeys *Cercopithecus diana*<sup>45</sup>).

It is worth noting that some conversational characteristics have also been found in nonprimate species. However, only the vocal interactions of nonhuman primates match all the criteria defining a “conversation-like vocal exchange” (see above the mentioned features). The origins of such vocal behaviours in nonhuman primates remain debated, notably because most published research on the topic has focused on monkeys, and additionally available findings in nonhuman apes (phylogenetically closer to humans) are controversial. The current state of knowledge prevents to conclude between a convergent evolution guided by the requirements of social life or a shared inheritance, suggesting an ancient mechanism which was already present in the primate lineage.

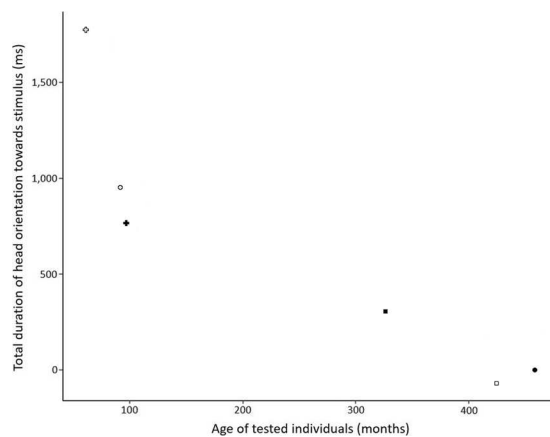
To our knowledge, three studies have focused on vocal conversational rules in nonhuman great apes, two of which found support for turn-taking and call overlap avoidance (bonobos<sup>40</sup>; gorillas<sup>43</sup>), and one that did not (chimpanzees<sup>46</sup>). Authors also questioned the fact that these abilities to converse in animals are genetically programmed or, as in humans<sup>47,48</sup>, socially acquired. Empirical evidence is currently lacking to answer this question, although the few available studies on monkeys showed that juveniles break conversational rules more often than do adults (howler monkeys *Alouatta pigra*<sup>21</sup>; Campbell’s monkeys<sup>49</sup>). Moreover, playback experiments using a violation of expectation paradigm showed that adult monkeys clearly discriminated between conversationally appropriate and inappropriate vocal exchange patterns, whereas socially inexperienced juveniles did not (Japanese macaques<sup>50</sup>; Campbell’s monkeys<sup>49</sup>). These experiments show that the audience detects the violation of expected social rules in vocal interactions between third parties<sup>49–52</sup>, suggesting that the social rules in vocal interactions are not a simple neurobiological guided behaviour but a social awareness of the respect of social rules.

Here, we aimed to progress this debate by conducting a playback experiment with captive western lowland gorillas of different ages. A recent study, ran on a captive group of western lowland gorillas housed at the ZooParc de Beauval, found that individuals do seem to engage in socially- and temporally-ruled vocal interactions<sup>43</sup>. Vocal interactions typically involve grunts, defined as soft contact calls<sup>53–55</sup> that are acoustically individually distinctive<sup>55</sup>. Grunts seem to play a key role in the coordination and peaceful activities, and thus act as a contact call<sup>53–55</sup>. During grunt exchanges, temporal rules are respected with an average “response” delay of 0.5 seconds (maximum 3 s) and with obvious call overlap avoidance<sup>43</sup>. Also, preferred interlocutors are non-randomly chosen. Grunt exchanges are mainly dyadic interactions that are most frequently observed among partners close in age, regardless of kinship<sup>43</sup>.

To assess the social relevance of these two potential conversational rules (inter-call durations and the age difference of interlocutors) for the audience in the captive population housed at the ZooParc de Beauval ( $n = 6$ ), we compared the responses of gorillas to grunt exchanges that have been built to match both the temporal and the social rules (condition A), in comparison with built grunt exchanges which respectively violate the overlap avoidance expectation (condition B) and the expectation of an age proximity between vocally exchanging partners (condition C). We also tested whether the age of the subject (perceiver) influenced its responses to these different vocal exchanges. We expected a modification of the attentional state of the audience by comparing the difference of the total duration of head orientation and the total number of occurrences of locomotion in the direction of the loudspeaker in the 60 seconds after the onset of the second vocalisation and 60 seconds before playback. We also examined the first latency to reposition the head front, corresponding to the duration of the first gaze towards the loudspeaker after the diffusion. The change of attentional state may be in both ways: an increased response towards the non-congruent pattern (implying that animals are surprised to hear the violation of the overlap avoidance and the expectation of an age proximity between interlocutors, as in Japanese macaques<sup>50</sup>), or towards the natural ones that they are exposed to in their everyday social life (implying that the audience is strongly interested by a relevant interaction including their congeners, as in Campbell’s monkeys<sup>49</sup>).



**Figure 1.** Total duration of head orientation toward the loudspeaker (in the 60 seconds ‘after’ minus 60 seconds ‘before’ playback) in the three playback conditions: “likely vocal exchange” (condition A), “overlapped vocal exchange” (condition B), “age-difference vocal exchange” (condition C). Each symbol corresponds to an individual. Box and whisker plots report the median, 25th and 75th percentiles, and the lowest and highest data, which are no more than 1.5 interquartile range from the box. \* $P < 0.05$ .

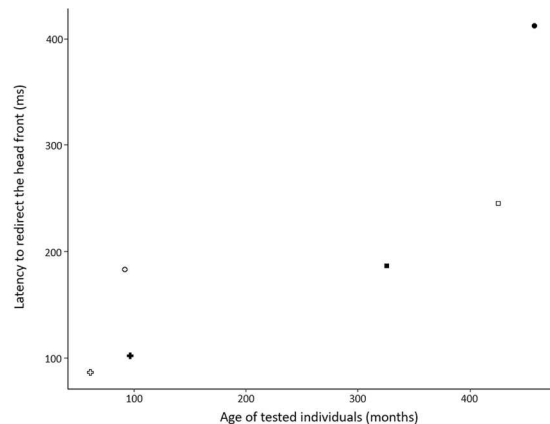


**Figure 2.** Total duration of head orientation in the “age-difference vocal exchange” (condition C) (in the 60 seconds ‘after’ minus 60 seconds ‘before’ playback) in relation to the subject’s age. Each symbol corresponds to an individual (same as in Fig. 1).

## Results

Analyses showed that individuals oriented their head towards the stimulus for a longer total time when temporal and social rules were respected (condition A) than when temporal rule was not (*i.e.* when calls were overlapping, condition B) (Wilcoxon signed rank test,  $V = 21$ ,  $P = 0.031$ , Fig. 1). However, there was no significant difference in the duration of head orientations between the likely (condition A) and the age-difference (condition C) vocal exchange conditions ( $V = 12$ ,  $P = 0.844$ ), nor between the overlapped calls and the age-difference vocal exchange conditions ( $V = 8$ ,  $P = 0.6875$ ). Moreover, no effect of condition was found for latency to redirect the head after the first gaze in direction to the loudspeaker (condition A: mean  $\pm$  SD =  $2025 \pm 1181$  msec; condition B:  $2005 \pm 72$  msec; condition C:  $4995 \pm 7256$  msec;  $V_{A/B} = 10$ ,  $P = 1$ ;  $V_{A/C} = 7$ ,  $P = 0.563$ ) nor for locomotion (condition A:  $0.22 \pm 0.29$ ; condition B:  $0.13 \pm 0.31$ ; condition C:  $0.071 \pm 0.21$ ;  $V_{A/B} = 5$ ,  $P = 0.423$ ;  $V_{A/C} = 6$ ,  $P = 0.174$ ).

The behavioural responses of the subject differed in some conditions according to its age. In the “age-difference vocal exchange” (condition C), we found that the older the subject, the shorter the head orientation duration towards the speaker (Spearman test,  $r_s = -0.94$ ,  $P = 0.017$ , Fig. 2). The total head orientation duration did not vary with subject’s age in the two other conditions ( $r_{sA} = 0.31$ ,  $P = 0.564$ ;  $r_{sB} = 0.03$ ,  $P = 1$ ). Likewise, when focusing on the latency to redirect the head after the first gaze directed to the speaker, we found that the older the subject, the longer the first head orientation in the “likely vocal exchange” (condition A) ( $r_s = 0.94$ ,  $P = 0.017$ , Fig. 3), but not in the two other conditions ( $r_{sB} = 0.49$ ,  $r_{sC} = -0.49$ ,  $P = 0.356$  in both cases). However, the age of the subject did not predict locomotion ( $r_{sA} = -0.76$ ,  $P = 0.08$ ,  $r_{sB} = -0.13$ ,  $P = 0.805$ ,  $r_{sC} = 0.34$ ,  $P = 0.512$ ).



**Figure 3.** Latency to redirect the head front in the “likely vocal exchange” (condition A) in relation to the subject’s age. Each symbol corresponds to an individual (same as in Fig. 1).

## Discussion

The violation-of-expectation paradigm is typically used in research on human and nonhuman animals to evaluate their ability to perceive certain rules<sup>56</sup>. Duration of post-playback head orientation responses is a common measure to assess a subject’s hearing general interest in a presented stimulus among both nonhuman<sup>57,58</sup> and human<sup>59</sup>, notably newborn<sup>60</sup> primates. In our study, results from the total duration of head orientation suggest that the gorilla’s audience pays attention to the vocal interactions among its group members and is more attentive to a “likely-vocal exchange” (condition A) than to an “overlapped vocal exchange” (condition B). In such “violation of attempt” setup, it is often predicted that unexpected stimuli would cause stronger responses than expected ones<sup>61</sup>. Indeed, nonhuman primates have shown more head orientations in response to a violation of the original pattern in several playback experiments<sup>62,63</sup>. Although our results may appear surprising in light of such studies, numerous authors have also reported an opposite effect of attempt violation on subjects’ responses. For instance, human infants tested in cross-modal studies show a familiarity/consistency preference<sup>64,65</sup>. The same is true for nonhuman primates. When exposing Japanese macaques to videos of conspecific facial expressions during vocal emissions, alongside the playback of a vocalisation which only matches one of the videos, subjects preferentially look at the congruent one (*i.e.* when vocalisation and video matched)<sup>66</sup>. A stronger attentional state towards an expected vocal sequence pattern rather than an unexpected one was also found in common marmosets<sup>67</sup>. Thus, the direction of the subjects’ attention bias in such experimental paradigms may be stimulus- or context-dependent. Interestingly, Lemasson and colleagues, who aimed answering the same research question as ours, obtained the same pattern in the Campbell monkeys they tested<sup>49</sup>.

In our study, the fact that the subjects paid a differential attention to the different test conditions (namely a strongest attention towards the “likely-vocal exchange” than the “overlapped vocal exchange”) suggests that the pattern was perceived and that temporal rules could be relevant to them. In the “overlapped vocal exchanges”, gorillas may not be able to extract all the relevant information, and notably not easily identify the callers, which can explain the weak responses of the audience in this condition. This can also be due to the fact that individuals evaluated the opportunity to participate in the ongoing simulated vocal exchange and lost interest quickly when they realised that the exchange was inappropriate, as suggested by Lemasson *et al.*<sup>49</sup> after similar results obtained in adult Campbell’s monkeys. We cannot totally rule out other possible alternative interpretations (although not mutually exclusive), notably that the lack of attention towards the “overlapped vocal exchange” (condition B) could be due to the fact that this stimulus is shorter in duration (so with a smaller quantity of “acoustic” information). However, the short duration hypothesis is unlikely for two reasons. First, because the difference in mean duration of the stimuli respecting temporal rules (*i.e.* conditions A and C) and those not respecting them (*i.e.* condition B) is very small (see methods). Second, because we found no significant difference between conditions B and C, although their stimuli differ in duration. This also reinforces our idea that the possible message confusion associated with call overlapping is not the core origin of the attentional bias observed above. The avoidance of call overlap could be a relevant social rule in the interactions of western gorillas, but did not elicit an attentional response strong enough to induce locomotion toward the loudspeaker. This is also true for the other conditions and not surprising given the fact that the simulated vocal exchanges involved soft contact calls emitted during peaceful activities<sup>53–55</sup>. Anyhow, such importance of the temporal organization of vocal interactions has already been highlighted in several monkey species<sup>25,49</sup>. In Campbell’s monkeys and Japanese macaques respectively, individuals that violate the turn-taking rule are typically young and inexperienced juveniles<sup>49</sup> or low-ranking (socially isolated) adult males<sup>68</sup>. In most human cultures, conversational overlap denotes a conversation failure and can lead to the end of an exchange<sup>3</sup>. In fact, vocal overlapping is perceived as a serious impoliteness in many traditional human societies<sup>69</sup>. In birds as well, overlap increases with aggressiveness and may lead to the end of an exchange (for a review see<sup>4</sup>). For example, in black capped chickadees (*Poecile atricapillus*), dominant males overlap more than do subordinates<sup>70,71</sup> and this overlapping increases arousal in overlapped individuals<sup>72,73</sup>.

The importance of age proximity between interlocutors for a gorilla audience remains unclear. Indeed, the age proximity or the age distance between interlocutors did not elicit different responses at the group level. One possible explanation is that breaking the social (age difference) rule is twice more common (15%) than breaking the

temporal (overlap) rule (8%) in gorillas and thus less surprising. The systematic focus on the silverback adult male in all broadcast vocal exchanges could have however induced a bias of general interest (*i.e.* silverback male vocalizations would systematically attract the attention of his groupmates whichever the context they are produced in), given that the role of the adult male is central in the social cohesion and coordination of the group<sup>74,75</sup>. This dominant male indeed develops tight bonds with each individual and is a key interlocutor for all group members. This supposed social rule (based on observations of naturally produced vocal exchanges<sup>43</sup>) cannot be considered as mandatory at this stage given that it does not elicit an unanimous behavioural response to violation.

Interestingly, the behavioural responses of the audience varied between individuals according to their age. The total duration of head orientation shows that the stronger attentional state of young gorillas toward the age-difference vocal interaction (*i.e.* condition C) is not trivial, and could reveal their interest for an eventual interaction involving another young individual and the adult male. That could explain why we did not find similar correlations in “likely vocal exchange” and “overlapped vocal exchange” conditions (*i.e.* conditions A and B) involving only adult individuals. This result suggests that, even if not a mandatory rule, gorillas may have expectations in this regard. Moreover, when we considered the first latency to reposition the head front as an indicator of the level of interest towards the stimulation, older gorillas paid more attention to exchanges between interlocutors close in age respecting common response delay (condition A) than did younger gorillas. Older gorillas may be socially more experienced than younger ones and more reliably perceive if a vocal interaction is relevant, as the “likely vocal exchange (*i.e.* condition A), or not, as the “overlapped vocal exchange” and the “age-difference vocal exchange” (*i.e.* conditions B and C), explaining the absence of correlation in conditions B and C. Nonetheless, these findings support the prediction, earlier suggested in monkeys<sup>49</sup>, that appropriate rules of conversing may be socially learned. Differences between adult and young individuals regarding the appropriateness of the context of call emission have already been reported in various primate species (see food calling in tamarins *Saguinus oedipus*<sup>76</sup>, and alarm calling in vervet monkeys *Chlorocebus pygerythrus*<sup>77</sup>). Some studies have also showed vocal development in the ability to interact properly with congeners (see vocal synchronisation in gibbons *Hylobates agilis agilis*<sup>78</sup>, as well as call matching and choice of preferred interlocutors in spider-monkeys<sup>38</sup>). Moreover, two playback studies have confirmed that conversational rules can be detected by adults and not by juveniles (turn-taking in Campbell’s monkeys<sup>49</sup>, and call matching in Japanese macaques<sup>50</sup>).

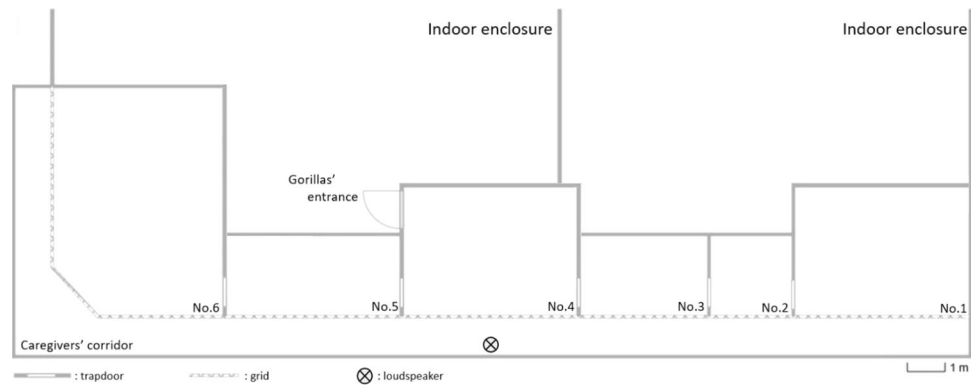
We acknowledge that our study is just one step towards a more advanced understanding of vocal interactions in great apes, and given the small sample size, some factors other than age might play a role. Indeed, despite all the advantages that the captivity offers to the study of vocal communication, gorillas cannot be isolated during experiments because of the risk of intense stress. Despite the randomisation of subject and condition orders and despite the fact that playbacks were time-spaced (over two months), a habituation effect cannot be discarded. Lastly, our result showing that young individuals responded strongly towards the “age-difference situation” (*i.e.* condition C) may suggest that the attentional response between each condition was driven by the age of individuals. Then, this difference of attentional state of the youngsters towards simulated vocal exchanges involving individuals of their own age range could induce a bias, mainly in the “overlapped vocal exchange” condition which include only adult individuals. To address this possible bias, a supplementary condition may be considered in future comparable studies. It would be similar to the “overlapped vocal exchange” but involving a young individual as first emitter who vocally interacts with the adult male, as the simulated vocal exchanges played back in this study (condition C). Our results are nevertheless promising and seem pointing out the possible existence of conversational rules managing vocal interactions in western lowland gorillas<sup>43</sup>. Future studies may consider investigating multiple and larger groups with a greater age-category range, and may test for potential intergroup variations. In the same vein, Levréro and colleagues<sup>40</sup> highlighted that bonobos are able to spontaneously display primitive conversation rules guided by social bonds, and Lameira and colleagues<sup>79,80</sup> showed that orang-utans (*Pongo pygmaeus*) have the cognitive abilities to produce spontaneous vocal exchanges. Only one study in apes, conducted in the field on male chimpanzees, revealed that the production of long calls (*i.e.* pant-hoots) is non-temporally organized (*i.e.* emissions are rather isolated calls without consecutive calls from other group members, or cases of synchronous choruses<sup>46</sup>). A recent study however has showed that chimpanzees have the ability to engage in gestural turn-taking<sup>81</sup>. All these findings suggest that all great apes possess the cognitive skills to engage in a “conversation-like” interaction. These results are promising and deserve further comparative work.

In conclusion, these recent studies on great apes highlight how vocal interactions are governed by social rules in our closest relatives, as they are across human cultures. Our findings help to fill the gap between nonhuman and human primate communication, particularly in terms of our understanding of factors that shape vocal interactions. The increase of the complexity of social interactions (from birds to mammals including monkeys, nonhuman apes and humans) may be at the origin of the most complex rules governing human communication. Parallels between human language and nonhuman primate vocal communication appear multifaceted, *e.g.*, information coding in acoustic structure, syntactic structure and socially-determined plasticity<sup>49,82,83</sup>. Here, a vast research area is opening into the developmental acquisition of conversational rules. The function, as well as the content of information in vocal exchanges remains unclear and still needs to be explored more systematically. Yet what is clear, is that the capacity to interact vocally with others and to respect ruled vocal interactions could have been an important step in the evolution of the vocal communication of primates<sup>36</sup>. This ability appeared before the capacity to articulate sounds in speech and to learn new acoustic structures, and is thus likely rooted deep in the animal lineage.

## Methods

**Compliance with ethical standards.** All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. The study has been conducted in accordance with the current laws in France (agreement with the 2010/63/UE). Animal husbandry and care were under the management of the animal caregivers of the ZooParc de Beauval, France. Our study was based on non-invasive observations and





**Figure 4.** Diagram of the layout of the cages.

Subject name (birth date)	Age-sex category <sup>a</sup>	Mother's name (all fathered by the adult male)	Cage number (date and new cage number if changed during the experimental period)	Number of calls used in built vocal exchanges
Inge (02/03/1980)	Adult female		No. 1	4
Kabinda (10/12/1982)	Adult female		No. 5	4
Sheila (18/02/1991)	Adult female		No. 6	4
Mapenzi (14/04/2010)	Subadult male	Kabinda	No. 3	0
Kuimba (19/08/2010)	Subadult female	Tamarilla <sup>†</sup>	No. 2	2
Mayélé (21/03/2013)	Juvenile female	Kabinda	No. 5 (05/11/2018, No. 3)	2

**Table 1.** Tested individuals in the payback experiment. <sup>a</sup>Age-sex categories defined according to Gatti and colleagues<sup>89</sup>; <sup>b</sup>Loudspeaker location; <sup>†</sup>Dead individual (10/02/2017).

was approved by the ethics committee “Comité Rennais d’Ethique en matière d’Expérimentation Animale” (*i.e.* Rennes Ethical comity for experiments using animals; CNREEA 2018041710224608-APAFIS#14920).

**Animals and housing conditions.** The sample comprised twelve captive-born western lowland gorillas, including one adult silverback male, four adult females, and seven (four males and three females) immature individuals. The composition of the captive group reflects that of wild groups<sup>84,85</sup>. All adult females were unrelated. The adult male had sired all the immature individuals of the group.

The animals were housed at the ZooParc de Beauval (France) in an enriched building of 720 m<sup>2</sup> (6 m high, with glass walls, wired roof, straw litter and wood perches, as well as temperature and humidity control), connected to several adjacent cages totalling 59 m<sup>2</sup> (average size: 5 m<sup>2</sup>) with visual and auditory access between them. Cages were aligned from right to left (*i.e.* from No. 1 to No. 6), and cage No. 4, of the silverback male, was central (Fig. 4). Depending on the weather conditions, group had access to a 3,500 m<sup>2</sup> outdoor enclosure (with grass, bush, and wood perches). Gorillas shared the enclosure with eight patas monkeys (*Erythrocebus patas*).

Gorillas were trained to enter the cages daily for breakfast and medical-training between 8:30 and 10:30 a.m., returning afterward to the large enclosure in which they spent the rest of the day and night. While they were free to go in any cages, each gorilla used to settle every day in a determined cage (one individual per cage, except for mother and offspring who may stay together). Once they were in their usual preferred cage, caregivers closed the trapdoors between each cage in order to avoid food theft (see Table 1 for cage occupation).

**Vocal recordings.** Between April and July 2018, we recorded 474 spontaneous grunts, produced by the twelve individuals during 33 breakfast sessions for a total of 27 hours of recording, using all occurrence sampling<sup>86</sup>. We pooled together all possible grunt structures (*i.e.* atonal and tonal grunts, single and double grunts<sup>53,54</sup>). Recordings were done at an average distance of 2 meters using a Marantz PMD660 handheld digital audio recorder (Marantz, Japan) (sample rate 44.1 kHz, resolution 32bits, wav. format) connected to a directional Sennheiser MKH70-1 microphone (Sennheiser, Germany). We identified 125 vocal interactions, based on Lemasson and colleagues’ definition<sup>43</sup> for gorillas (*i.e.* series of calls emitted by different callers with an inter-call duration of less than 3 seconds). Vocal interactions typically involved callers from different cages (usually only two callers: 77%) and non-overlapped exchanged calls (92%). The adult male contributed 70% of these vocal interactions and was significantly more often a respondent than an initiator (Binomial test,  $N_{1st\ position} = 20$ ,  $N_{2nd\ position} = 41$ ,  $p = 0.01$ ). He responded more often to adult females (85% of response) than to young individuals (*i.e.* sub-adults, juvenile and infants, 15%).

**Playback experiments.** Playback experiments took place between September and December 2018. Participating gorillas never experienced any playback experiment in the past. Here, we tested their responses to

the broadcast of three types of vocal exchanges, *i.e.* one control condition, one violating the temporal expectation and one violating the social expectation:

- Condition A, “likely vocal exchange”: we built a control dyadic vocal exchange (as described in Lemasson *et al.*<sup>41</sup>) between two individuals close in age (*i.e.* both adults), that respected the average inter-call duration of 500 milliseconds (ms) (see Supplementary Fig. S1);
- Condition B, “overlapped vocal exchange”: we built an unlikely vocal exchange that did not obey the temporal turn-taking rule, with 50% of overlap between two vocalisations<sup>87</sup> emitted by two adults (see Supplementary Fig. S2);
- Condition C, “age-difference vocal exchange”: we built an unlikely vocal exchange, with an inter-call duration of 500 ms, that did not obey the social rule with two callers of distant ages (*i.e.* initiator: sub-adult or juvenile, mean age  $\pm$  SD =  $6.39 \pm 1.33$  years; respondent: adult male, 27 years; see Supplementary Fig. S3).

Vocal exchange acoustic stimuli were created using the vocalisations recorded during the aforementioned observations, discarding non-exchanged grunts, and using PRAAT and Goldwave softwares. Calls of initiators and respondents were systematically concatenated to respect their original position from dyadic natural vocal interactions previously recorded. Even the control dyadic “likely vocal exchanges” were built to prevent any bias due to artificial concatenation of calls. To match the most frequent vocal events occurring in natural conditions, we decided to use the adult male as the respondent in all acoustic stimuli and different females as initiators. To prevent kin-bias, subjects only heard unrelated female initiators. Thus, a given adult female gorilla was allotted to a given subject in conditions A and B. To prevent pseudo-replication, a given call exemplar was used only once. Given that spontaneous vocal interactions (*i.e.* abovementioned recordings) contain an equal number of grunts and double grunts, we used 50% of grunts and 50% of double grunts. The total durations of the played back stimuli in conditions A and C (excluding inter-call durations) were respectively  $163 \pm 26$  ms and  $146 \pm 30$  ms, and  $125 \pm 36$  ms in condition B.

Experiments took place during the usual breakfast in the corridor facing the cages. Sounds were played using a TAG Premio8.2 loudspeaker connected to the Marantz player at 60 dB SPL. The intensity of 60 dB corresponds to the grunts recorded in this group in similar conditions at an average distance of 2 meters. This intensity threshold matches the intensity used in other nonhuman primate studies (*e.g.* Japanese macaques<sup>50</sup>, bonobos<sup>88</sup>). The loudspeaker was systematically placed before the arrival of gorillas in the cages at 1 meter in front of cage No.4 (of the silverback male, see Fig. 4) and could not be seen by the other gorillas. The experimenter (LP) sat in front of the tested animal and fed it routinely with fruits, dry-food and lastly vegetables (always in this order), through the wire mesh. The camera was positioned behind the experimenter, above his head. The breakfast typically lasted around 30 minutes and the playback occurred at mid-term. The experimenter offered food to the subject in order to control the central position of his body and head at T0 (*i.e.* playback time). Before broadcasting the stimuli, some conditions had been respected to make the situation as plausible as possible: (i) both initiators and respondents used in a stimulated vocal exchange on the same side of the loudspeaker, both at the opposite side of the tested individual, and were not visible from the subject point of view, (ii) no vocalisation was heard within the two minutes preceding the playback. To prevent any possible influence from the experimenter, the experimenter’s gaze had to focus on the gorilla’s chest during the entire feeding period. To habituate animals to the experimental setting, we ran ‘silent sessions’ during which the full equipment was set up to mimic the conditions of real playback sessions but without any sound broadcast. Twenty days before starting the first playback session, we performed ‘silent sessions’ once a day, every day. We then performed randomly ‘silent sessions’ throughout the entire duration of the experiment. To limit habituation, no more than one subject was tested per day and each subject was tested at average intervals of 27 days (min: 3, max: 73 days apart). This variation is due to the fact that we had to stop our experiments on two occasions, after 11 and 28 days respectively, because of on-going conflicts in the group or construction work in the enclosure.

Six individuals were tested in the playback paradigm (Table 1). The three infants were excluded from the study as they were still dependent on their mothers. The silverback male was not tested as he was involved in the creation of all acoustic stimuli. We discarded one adult female and one subadult female, because they were too stressed during the experiments, either by the presence of the experimenter or by a conflict with other group members. Each tested gorilla heard each of three distinct conditions, resulting in 18 (3  $\times$  6) stimuli in total. Subject order and condition order were both randomized.

**Recording of behavioural responses.** We filmed the tested gorilla’s behaviours before and after playback. All videos were analysed in slow motion (scale of millisecond) using Kinovea 80.8.15 software for coding, while the experimenter was ‘blind’ to the experimental condition. The behavioural responses measured were the difference in the total duration of head orientation and in the total number of occurrences of locomotion (already observed in bonobos towards similar stimuli<sup>87</sup>) in the direction of the loudspeaker before and after the playback. The time windows used for the measurements were 60 seconds before and 60 seconds after the onset of the second vocalisation (*i.e.* T0 = onset of male calling, as this is the moment when subjects typically detect the breaking of the rule, as in Bouchet and colleagues<sup>50</sup>). This correction was chosen because contact call exchanges are not rare events that would typically trigger a strong surprise, because inter-cage gaze exchanges are frequent at breakfast time and lastly because subjects differed in their baseline interest towards their social environment; we thus wanted to capture not only the degree of interest towards the playback area but more importantly the increase/decrease of interest towards this area, as in other previous studies<sup>57,90,91</sup>. We also measured the first latency to reposition the head front (*i.e.* facing front the experimenter), corresponding to the duration of the first gaze towards the loudspeaker (starting the measure at T0 if the individual turned its head before the onset of the second vocalisation).

Normality and homogeneity of variance were assessed by inspecting residuals following Shapiro-Wilk *W* tests. Given the non-normally distributed data, we ran non-parametric tests<sup>92</sup> with a significance threshold of 0.05. The difference of total duration of head orientation, of locomotion occurrences, and the latencies to reposition the head between likely (condition A) and unlikely vocal exchanges (conditions B and C) were tested using exact Wilcoxon signed-rank tests. Spearman's correlation tests were then run to investigate the possible influence of the age of the subjects on their responses to each situation. Data were analysed in R Studio0.99.903 (R version3.3.1: R Core Team2016).

## Data availability

The data generated and analysed during the current study are available from the corresponding author on reasonable request.

Received: 23 September 2019; Accepted: 7 April 2020;

Published online: 24 April 2020

## References

1. Stivers, T. *et al.* Universals and cultural variation in turn-taking in conversation. *Proc. Natl. Acad. Sci.* **106**, 10587–10592 (2009).
2. Ghiglione, R. *L'homme communiquant.* (1986).
3. Sacks, H., Schegloff, E. A. & Jefferson, G. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language (Baltim)*. **50**, 696–735 (1974).
4. Henry, L., Craig, A. J. F. K., Lemasson, A. & Hausberger, M. Social coordination in animal vocal interactions. Is there any evidence of turn-taking? The starling as an animal model. *Front. Psychol.* **6**, 1416 (2015).
5. Lemasson, A., Ouattara, K., Petit, E. J. & Zuberbühler, K. Social learning of vocal structure in a nonhuman primate? *BMC Evol. Biol.* **11**, 1–7 (2011).
6. Snowdon, C. T. & Cleveland, J. "Conversations" among pygmy marmosets. *Am. J. Primatol.* **7**, 15–20 (1984).
7. Sugiura, H. & Masataka, N. Temporal and acoustic flexibility in vocal exchanges of coo calls in Japanese macaques (*Macaca fuscata*). in *Current topics in primate vocal communication* (eds. Zimmermann, E., Newman, J. D. & Jürgens, U.) 121–140, <https://doi.org/10.1007/978-1-4757-9930-9> (1995).
8. Bouchet, H., Pellier, A. S., Blois-Heulin, C. & Lemasson, A. Sex differences in the vocal repertoire of adult red-capped mangabeys (*Cercocebus torquatus*): A multi-level acoustic analysis. *Am. J. Primatol.* **72**, 360–375 (2010).
9. Koda, H. Flexibility and context-sensitivity during the vocal exchange of coo calls in wild Japanese macaques (*Macaca fuscata yakui*). *Behaviour* **141**, 1279–1296 (2004).
10. Papworth, S., Böse, A.-S., Barker, J., Schel, A.-M. & Zuberbühler, K. Male blue monkeys alarm call in response to danger experienced by others. *Biol. Lett.* **4**, 472–475 (2008).
11. Clark, A. P. & Wrangham, R. W. Acoustic analysis of wild chimpanzee pant hoots: Do Kibale Forest chimpanzees have an acoustically distinct food arrival pant hoot? *Am. J. Primatol.*, <https://doi.org/10.1002/ajp.1350310203> (1993).
12. Dahlin, C. R. & Benedict, L. Angry birds need not apply: A perspective on the flexible form and multifunctionality of avian vocal duets. *Ethology* **120**, 1–10 (2014).
13. Hall, M. L. A Review of Vocal Duetting in Birds. *Adv. Study Behav.* **40**, 67–121 (2009).
14. Hall, M. L. A review of hypotheses for the functions of avian duetting. *Behav. Ecol. Sociobiol.* **55**, 415–430 (2004).
15. Geissmann, T. & Orgeldinger, M. The relationship between duet songs and pair bonds in siamangs, *Hylobates syndactylus*. *Anim. Behav.* **60**, 805–809 (2000).
16. Forester, D. C. & Harrison, W. K. The Significance of Persistent Vocalisation By the Spring Peeper, *Pseudacris Crucifer* (Anura: Hylidae). *Behaviour* **103**, 1–15 (2008).
17. Bailey, W. J. Insect duets: Underlying mechanisms and their evolution. *Physiol. Entomol.* **28**, 157–174 (2003).
18. Okobi, D. E. J., Banerjee, A., Matheson, A. M. M., Phelps, S. M. & Long, M. A. Motor cortical control of vocal interaction in neotropical singing mice. *Sci.* **363**, 979–983 (2019).
19. Clutton-Brock, T. H. & Albon, S. D. The Roaring of Red Deer and the Evolution of Honest Advertisement. *Behaviour* **69**, 145–170 (1979).
20. Reby, D. & McComb, K. Vocal Communication and Reproduction in Deer. *Adv. Study Behav.* **33**, 231–264 (2003).
21. Briseño Jaramillo, M., Guyot-Biquand, V., Estrada, A. & Lemasson, A. Vocal repertoire of free-ranging black howler monkeys' (*Alouatta pigra*): call types, contexts and sex-related contributions. *Am. J. Primatol.* **79**, 22630–10 (2017).
22. Charlton, B. D. *et al.* Perception of male caller identity in koalas (*Phascolarctos cinereus*): Acoustic analysis and playback experiments. *Plos One* **6**, e20329 (2011).
23. Gros-Louis, J. & Mitani, J. Chorusing and call convergence in chimpanzees: Tests of three hypotheses. *Behaviour* **135**, 1041–1064 (1998).
24. France, E. F., Anderson, A. H. & Gardner, M. The impact of status and audio conferencing technology on business meetings. *Int. J. Hum. Comput. Stud.* **54**, 857–876 (2001).
25. Pika, S., Wilkinson, R., Kendrick, K. H. & Vernes, S. C. Taking turns: bridging the gap between human and animal communication. *Proc. R. Soc. B* **285**, 1–9 (2018).
26. Leong, K. M., Ortolani, A., Burks, K. D., Melleni, J. D. & Savage, A. Quantifying Acoustic and Temporal Characteristics of Vocalizations for a Group of Captive African Elephants *Loxodonta Africana*. *Bioacoustics Int. J. Anim. Sound its Rec.* **13**, 213–231 (2003).
27. Schulz, T. M., Whitehead, H., Gero, S. & Rendell, L. Overlapping and matching of codas in vocal interactions between sperm whales: insights into communication function. *Anim. Behav.* **76**, 1977–1988 (2008).
28. Janik, V. M. Whistle Matching in Wild Bottlenose Dolphins (*Tursiops truncatus*). *Science (80-)*. **289**, 1355–1357 (2000).
29. Yoshida, S. & Okanoya, K. Evolution of turntaking: a bio-cognitive perspective. *Cogn. Stud.* **12**, 153–165 (2005).
30. Carter, G. G., Skowronski, M. D., Faure, P. A. & Fenton, B. Antiphonal calling allows individual discrimination in white-winged vampire bats. *Anim. Behav.* **76**, 1343–1355 (2008).
31. Lemasson, A., Gandon, E. & Hausberger, M. Attention to elders' voice in non-human primates. *Biol. Lett.* **6**, 325–328 (2010).
32. Takahashi, D. Y., Narayanan, D. Z. & Ghazanfar, A. A. Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys. *Curr. Biol.* **23**, 2162–2168 (2013).
33. Kondo, N., Watanabe, S. & Izawa, E. I. A temporal rule in vocal exchange among large-billed crows *Corvus macrorhynchos* in Japan. *Ornithol. Sci.* **9**, 83–91 (2010).
34. Reber, S. A., Townsend, S. W. & Manser, M. B. Social monitoring via close calls in meerkats. *Proc. R. Soc. B Biol. Sci.* **280** (2013).
35. Soltis, J., Leong, K. & Savage, A. African elephant vocal communication I: Antiphonal calling behaviour among affiliated females. *Anim. Behav.* **70**, 579–587 (2005).
36. Dunbar, R. *Grooming, gossip and the evolution of language.* (Faber & Faber, 1996).



37. Cutting, A. L. & Dunn, J. Conversations with siblings and with friends: Links between relationship quality and social understanding. *Br. J. Dev. Psychol.* **24**, 73–87 (2006).
38. Briseno-Jaramillo, M. *et al.* Age and social affinity effects on contact call interactions in free-ranging spider monkeys. *Behav. Ecol. Sociobiol.* **72**, 192 (2018).
39. Arlet, M., Jubin, R., Masataka, N. & Lemasson, A. Grooming-at-a-distance by exchanging calls in non-human primates. *Biol. Lett.* **11**, 20150711–20150714 (2015).
40. Levréro, F. *et al.* Social bonding drives vocal exchanges in Bonobos. *Sci. Rep.* **9**, 711 (2019).
41. Bascom, W. R. The principle of seniority in the social structure of the Yoruba. *Am. Anthropol.* **44**, 37–46 (1942).
42. Chen, H. C., Kaplan, G. & Rogers, L. J. Contact calls of common marmosets (*Callithrix jacchus*): Influence of age of caller on antiphonal calling and other vocal responses. *Am. J. Primatol.* **71**, 165–170 (2009).
43. Lemasson, A., Pereira, H. & Levréro, F. Social Basis of Vocal Interactions in Western Lowland Gorillas (*Gorilla g. gorilla*). *J. Comp. Psychol.* **132**, 141–151 (2018).
44. Giles, H., Coupland, N. & Coupland, J. Accommodation Theory: Communication, Context, and Consequence. *Context. Accommod. Dev. Appl. Socioling.* 1–68 (1991).
45. Candiotti, A., Zuberbühler, K. & Lemasson, A. Convergence and divergence in Diana monkey vocalizations. *Biol. Lett.* **8**, 382–385 (2012).
46. Arcadi, A. C. Vocal responsiveness in male wild chimpanzees: Implications for the evolution of language. *J. Hum. Evol.* **39**, 205–223 (2000).
47. Ginsburg, G. P. & Kilbourne, B. K. Emergence of vocal alternation in mother-infant interchanges. *J. Child Lang.* **15**, 221–235 (1988).
48. Stern, D. N., Jaffe, J., Beebe, B. & Bennett, S. L. Vocalizing in unison and in alternation: Two modes of communication within the mother-infant dyad. *Ann. N. Y. Acad. Sci.* **263**, 89–100 (1975).
49. Lemasson, A. *et al.* Youngsters do not pay attention to conversational rules: Is this so for nonhuman primates? *Sci. Rep.* **1**, (2011).
50. Bouchet, H., Koda, H. & Lemasson, A. Age-dependent change in attention paid to vocal exchange rules in Japanese macaques. *Anim. Behav.* **129**, 81–92 (2017).
51. Cheney, D. L., Seyfarth, R. M. & Silk, J. B. The Responses of Female Baboons (*Papio cynocephalus ursinus*) to Anomalous Social Interactions: Evidence for Causal Reasoning? *J. Comp. Psychol.* **109**, 134–141 (1995).
52. Wittig, R. M., Crockford, C., Seyfarth, R. M. & Cheney, D. L. Vocal alliances in Chacma baboons (*Papio hamadryas ursinus*). *Behav. Ecol. Sociobiol.* **61**, 899–909 (2007).
53. Hedwig, D., Robbins, M. M., Mundry, R., Hammerschmidt, K. & Boesch, C. Acoustic structure and variation in mountain and western gorilla close calls: A syntactic approach. *Behaviour* **151**, 1091–1120 (2014).
54. Hedwig, D., Mundry, R., Robbins, M. M. & Boesch, C. Contextual correlates of syntactic variation in mountain and western gorilla close-distance vocalizations: Indications for lexical or phonological syntax? *Anim. Cogn.* **18**, 423–435 (2015).
55. Salmi, R., Hammerschmidt, K. & Doran-Sheehy, D. M. Individual distinctiveness in call types of wild western female gorillas. *PLoS One* **9**, e101940 (2014).
56. Zuberbühler, K. & Wittig, R. M. Field experiments with non-human primates: A tutorial. in *Field and laboratory methods in primatology: A practical guide* (eds. Setchell, J. M. & Curtis, D. J.) 207–224 (Cambridge University Press, 2011).
57. Lemasson, A., Palombit, R. A. & Jubin, R. Friendships between males and lactating females in a free-ranging group of olive baboons (*Papio hamadryas anubis*): Evidence from playback experiments. *Behav. Ecol. Sociobiol.* **62**, 1027–1035 (2008).
58. Levréro, F. *et al.* Social shaping of voices does not impair phenotype matching of kinship in mandrills. *Nat. Commun.* **6**, 7609 (2015).
59. Hayashi, A., Tamekawa, Y. & Kiritani, S. Developmental Change in Auditory Preferences for Speech Stimuli in Japanese Infants. *J. Speech, Lang. Hear. Res.* **44**, 1189–1200 (2001).
60. Kinzler, K. D., Dupoux, E. & Spelke, E. S. The native language of social cognition. *Proc. Natl. Acad. Sci.* **104**, 12577–12580 (2007).
61. Lewis, M. & Goldberg, S. The Acquisition and Violation of Expectancy: An Experimental Paradigm. *J. Exp. Child Psychol.* **7**, 70–80 (1969).
62. Saffran, J. *et al.* Grammatical pattern learning by human infants and cotton-top tamarin monkeys. *Cognition* **107**, 479–500 (2008).
63. Wilson, B. *et al.* Auditory artificial grammar learning in Macaque and Marmoset monkeys. *J. Neurosci.* **33**, 18825–18835 (2013).
64. Meltzoff, A. N. & Borton, R. W. Intermodal matching by human neonates. *Nature* **282**, 403–404 (1979).
65. Kuhl, P. K. & Meltzoff, A. N. The bimodal perception of speech in infancy. *Science* (80-). **218**, 1138–1141 (1982).
66. Ghazanfar, A. A. & Logothetis, N. K. Facial expressions linked to monkey calls. *Nature* **423**, 937–938 (2002).
67. Reber, S. A. *et al.* Common marmosets are sensitive to simple dependencies at variable distances in an artificial grammar. *Evol. Hum. Behav.* **40**, 214–221 (2019).
68. Lemasson, A. *et al.* Age- and sex-dependent contact call usage in Japanese macaques. *Primates* **54**, 283–291 (2013).
69. Calame-Griaule, G. *Ethnologie et langage. la parole chez les dogon.* (1965).
70. Baker, T. M., Wilson, D. R. & Mennill, D. J. Vocal signals predict attack during aggressive interactions in black-capped chickadees. *Anim. Behav.* **84**, 965–974 (2012).
71. Ficken, M. S., Ficken, R. W. & Witkin, S. R. Vocal Repertoire of the Black-Capped Chickadee. *Auk* **95**, 34–48 (1978).
72. Dabelsteen, T., McGregor, P. K., Holland, J., Tobias, J. A. & Pedersen, S. B. The signal function of overlapping singing in male robins. *Anim. Behav.* **53**, 249–256 (2003).
73. Mennill, D. J. & Ratcliffe, L. M. Overlapping and matching in the song contests of black-capped chickadees. *Anim. Behav.* **67**, 441–450 (2004).
74. Harcourt, A. H. Social relationships between adult male and female mountain gorillas in the wild. *Anim. Behav.* **27**, 325–342 (1979).
75. Watts, D. P. Gorilla social relationships: A comparative overview. in *Gorilla biology: A multidisciplinary perspective* 302–327 (Cambridge University Press, 2002).
76. Cheney, D. L. & Seyfarth, R. M. *How monkeys see the world.* (University of Chicago Press, 1990).
77. Seyfarth, R. M. & Cheney, D. L. Vocal development in vervet monkeys. *Anim. Behav.* **34**, 1640–1658 (1986).
78. Koda, H. *et al.* Possible Role of Mother-Daughter Vocal Interactions on the Development of Species-Specific Song in Gibbons. *PLoS One* **8** (2013).
79. Lameira, A. R., Maddieson, I. & Zuberbühler, K. Primate feedstock for the evolution of consonants. *Trends Cogn. Sci.* **18**, 60–62 (2014).
80. Lameira, A. R., Hardus, M. E., Mielke, A., Wich, S. A. & Shumaker, R. W. Vocal fold control beyond the species-specific repertoire in an orang-utan. *Sci. Rep.* **6**, 30315 (2016).
81. Fröhlich, M. *et al.* Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. *Sci. Rep.* **6**, 1–14 (2016).
82. Bouchet, H., Coye, C. & Lemasson, A. Le Langage est-il le Propre de l’Homme? Apports des Etudes sur les Primates Non Humains. *Tétralogiques* **21**, 87–133 (2016).
83. Zuberbühler, K. The phylogenetic roots of language: Evidence from primate communication and cognition. *Curr. Dir. Psychol. Sci.* **14**, 126–130 (2005).
84. Robbins, M. M. *et al.* Social structure and life-history patterns in western gorillas (*Gorilla gorilla gorilla*). *Am. J. Primatol.* **64**, 145–159 (2004).
85. Parnell, R. J. Group size and structure in western lowland gorillas (*Gorilla gorilla gorilla*) at Mbeli Bai, Republic of Congo. *Am. J. Primatol.* **56**, 193–206 (2002).

86. Altmann, J. Observational Study of Behavior: Sampling Methods. *Behaviour* **49**, 227–267 (1974).
87. Levréro, F. *et al.* Bonobos converse and pay attention to breaking conversational rules. in *Behaviour 2017 - 35th International Ethological Conference* (ed. Instituto Universitário de Ciências Psicológicas, S. e da V. (ISPA); A. for the S. of A. B. (ASAB)) hal-01576988 (2017).
88. Keenan, S. *et al.* Enduring voice recognition in bonobos. *Sci. Rep.* **6**, (2016).
89. Gatti, S., Levréro, F., Ménard, N. & Gautierhion, A. Population and group structure of western Lowland Gorillas (*Gorilla gorilla gorilla*) at Lokoué, Republic of Congo. *Am. J. Primatol.* **63**, 111–123 (2004).
90. Palombit, R. A., Seyfarth, R. M. & Cheney, D. L. The adaptive value of ‘friendships’ to female baboons: experimental and observational evidence. *Anim. Behav* **54**, 599–614 (1997).
91. Lemasson, A., Boutin, A., Boivin, S., Blois-Heulin, C. & Hausberger, M. Horse (*Equus caballus*) whinnies: a source of social information. *Anim. Cogn.* **12**, 693–704 (2009).
92. Siegel, S. & Castellan, N. J. *Nonparametric statistics for the behavioral sciences* (2nd ed.) (1988).

## Acknowledgements

We are deeply grateful to the ZooParc de Beauval and the gorilla caregivers for their valuable assistance. We also thank Katarzyna Pisanski and Melissa Sebilliau for correcting our English. This work was financially supported by a CIFRE grant (#2018\_00074) funded by the Association Beauval Nature pour la Conservation et la Recherche, Centre National de la Recherche Scientifique, Rennes 1 University and Lyon/Saint-Etienne University.

## Author contributions

L.P., F.L., B.M. and A.L. conceived, designed and coordinated the study. L.P. collected data and carried out the statistical analyses. L.P., F.L. and A.L. interpreted and drafted the manuscript. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-63923-7>.

**Correspondence** and requests for materials should be addressed to L.P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

# Breaking conversational rules matters to captive gorillas: a playback experiment

Loïc Pognault<sup>1,2,3</sup>, Florence Levréro<sup>2\*</sup>, Baptiste Mulot<sup>3</sup>, Alban Lemasson<sup>1\*</sup>

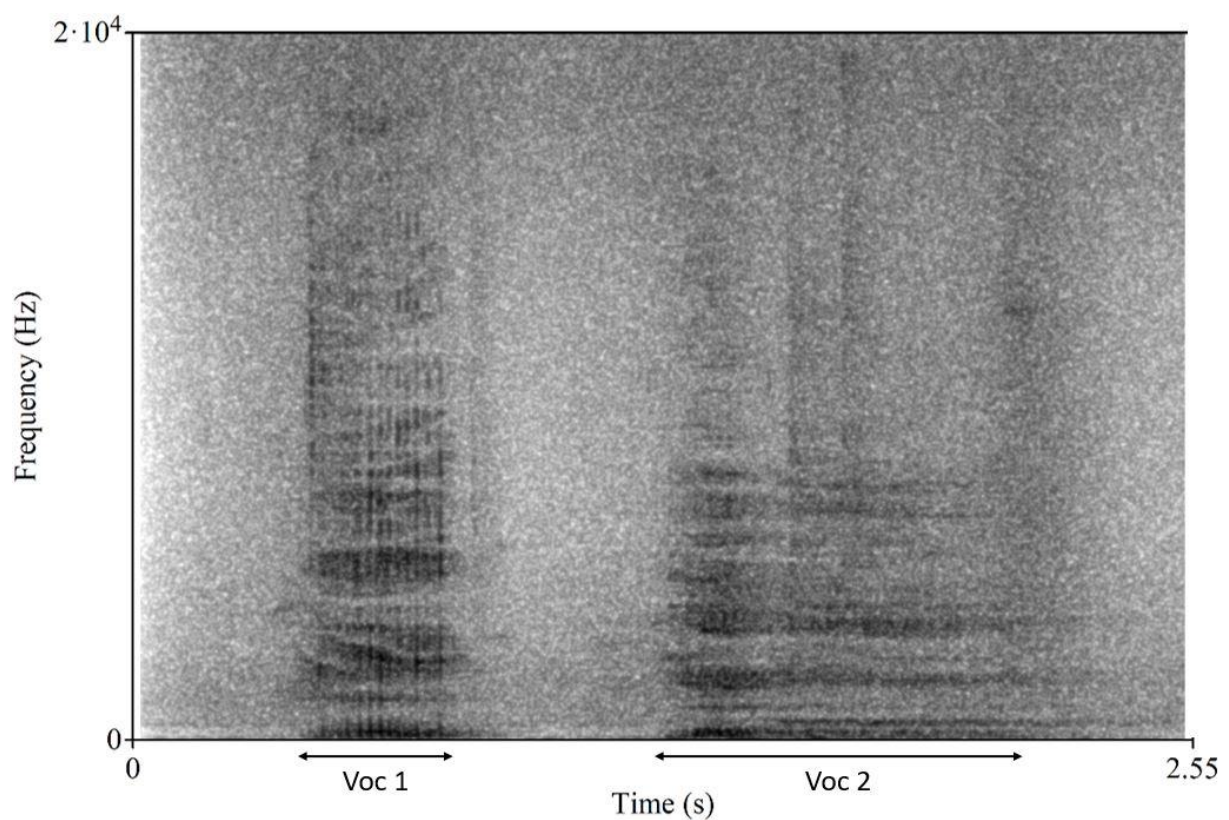
<sup>1</sup> Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000 Rennes, France.

<sup>2</sup> Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES / CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France.

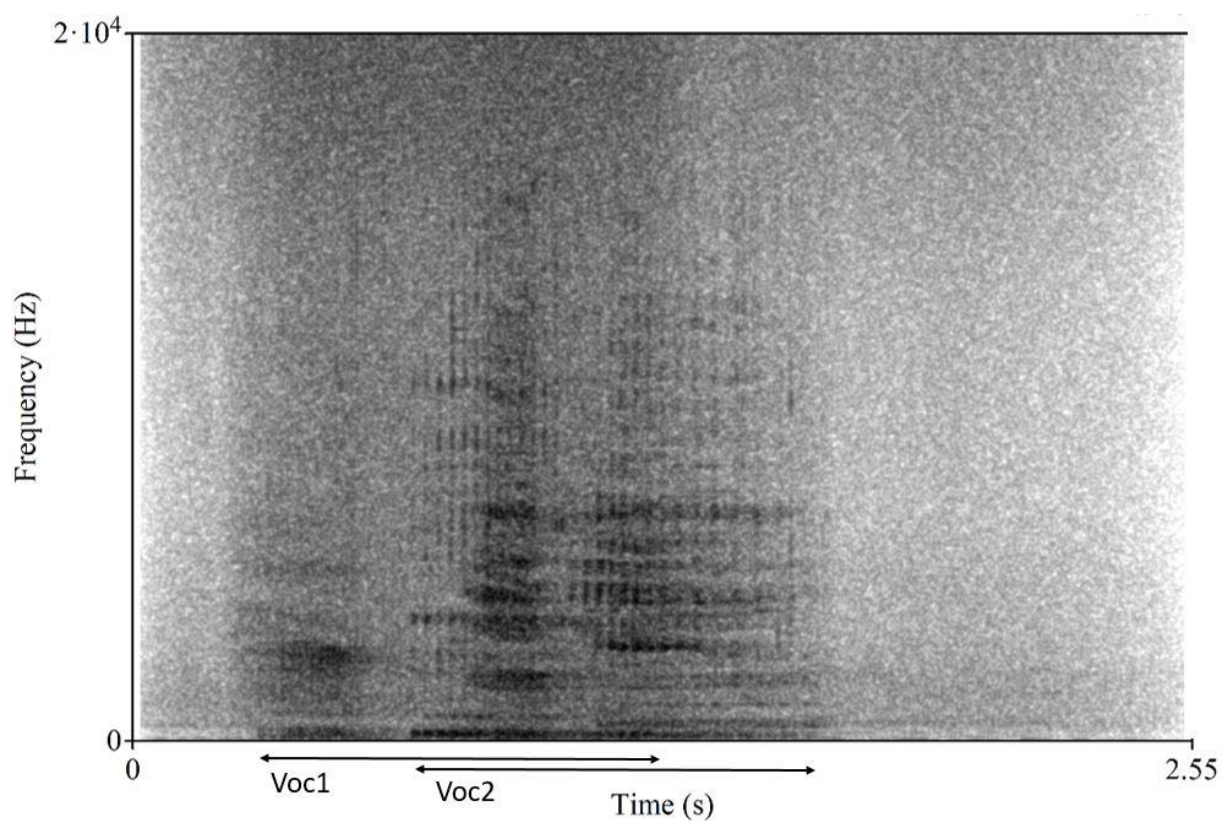
<sup>3</sup> ZooParc de Beauval & Beauval Nature, 41110 Saint Aignan, France

\* Same contribution

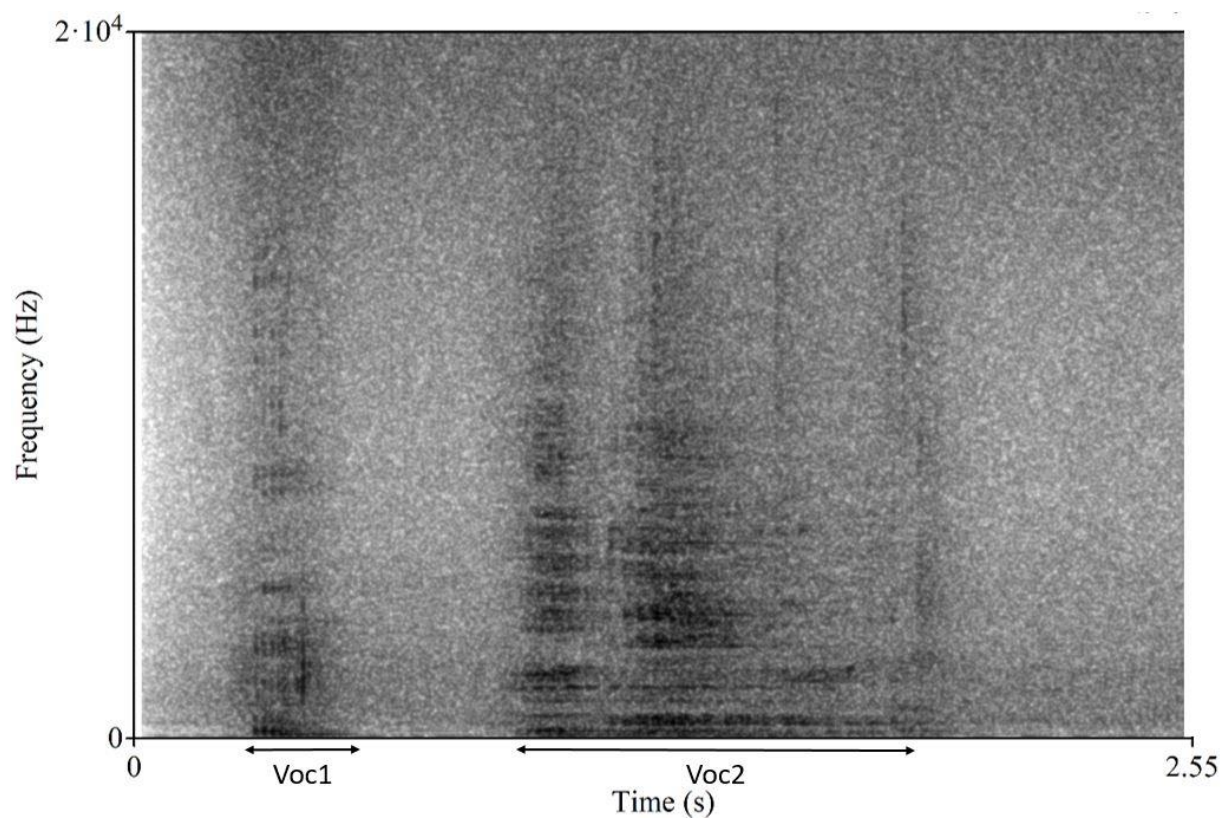
## Supplementary materials



**Figure S1** | Spectrogram of the “likely-vocal exchange”, Condition A



**Figure S2** | Spectrogram of the “overlapped vocal exchange”, Condition B



**Figure S3** | Spectrogram of the “age-difference vocal exchange”, Condition C



# Chapitre IV

## **Organisation temporelle des productions vocales et influences sociales**





# Article 3

## **Preliminary investigation of temporal calling patterns of a captive group of chimpanzees**

Loïc Pougnault<sup>1,2,3</sup>, Alban Lemasson<sup>1</sup>, Baptiste Mulo<sup>3</sup>, Florence Levréro<sup>2</sup>

<sup>1</sup> Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000 Rennes, France.

<sup>2</sup> Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES / CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France.

<sup>3</sup> ZooParc de Beauval & Beauval Nature, 41110 Saint Aignan, France

*Article en préparation*





## Synthèse : article 3



Les échanges vocaux organisés sur la base du tour de parole sont-ils réellement absents des interactions vocales des chimpanzés ?

### Contexte

En 2000, Arcadi conduit une première étude qui souligne l'absence d'échanges vocaux organisés chez les chimpanzés mâles de la communauté de Kibale (Ouganda), au profit de vocalisations isolées ou produites en *chorus*. A la vue des nombreuses découvertes chez les primates non-hominoïdes, et récemment chez certaines espèces de grands singes (*i.e.* gorilles et bonobos), ces résultats chez les chimpanzés sont étonnants. Par notre étude, nous avons cherché à réaliser une première description de l'organisation temporelle et sociale de l'ensemble des productions vocales chez les chimpanzés d'un groupe captif (*i.e.* mâles et femelles de tous âges), et ce en vue de confirmer ou non les précédents résultats d'Arcadi. Par ailleurs, nous avons testé l'impact que peuvent avoir les différentes organisations temporelles de productions vocales sur la réaction non-vocale de l'audience afin d'en évaluer la fonction sociale.

### Méthode

#### ① Observations & Enregistrements

##### Collectes des données sociales :

*Focal par individu (N<sub>ind</sub>=15)*  
1h30/ind. (Sessions de 10min/ind.)  
Total : 22h30

Relevés : Interactions sociales  
(Affiliatif/Agressif/de Soumission)

##### *Scans de groupe*

T= 10 min (Sessions de 1hr)  
Total : 108 scans sur le groupe

Relevés : Proximité interindividuelle

Indices considérés : Dominance, intégration Sociale & Spatiale +  
Age & Sexe

##### Collecte des données vocales :

###### *Toutes occurrences*

(Sessions de 1hr) | Contexte de la production + Réaction non-vocale de l'audience  
(rayon de 10/15m autour de l'émetteur pendant 10 sec)  
Total : 26hr30

#### ② Descriptions des productions vocales

##### Dimension temporelle

- A l'échelle du groupe :  
Emises en interaction VS hors interaction
- A l'échelle de l'individu :  
Emises : de façon unique VS répétée

##### Dimension sociale

- Influence des caractéristiques individuelles et contextuelles sur les différentes organisations temporelles de productions vocales.
- Influence des différentes organisations temporelles sur la réaction non-vocale de l'audience.

### Principaux résultats

- La vaste majorité des vocalisations sont produites hors interactions vocales, et les vocalisations superposées des mâles représentent le cœur des interactions vocales, alors que les échanges respectant le tour de parole sont extrêmement rares.
- Le contexte agressif favorise la production de vocalisations répétées, et les individus dominants les produisent plus fréquemment que le reste du groupe.
- Plus une vocalisation est répétée, et plus la probabilité d'observer une réaction non-vocale de l'audience est importante

### Conclusion

Selon notre étude, les vocalisations produites ne suscitent généralement pas la mise en place d'échanges vocaux organisés au sein du groupe, les interactions vocales se limitant pour l'essentiel à des vocalisations superposées des mâles. L'existence de *chorus* de vocalisations de forte intensité semble cohérente avec la socialité hautement compétitrice des chimpanzés, aussi bien au niveau intra qu'inter-communauté. Néanmoins, l'extrême rareté des échanges vocaux organisés que confirme notre étude, fait des chimpanzés une espèce atypique.

## **Preliminary investigation of temporal calling patterns of a captive group of chimpanzees**

**Loïc Pougault<sup>1,2,3</sup>, Alban Lemasson<sup>1\*</sup>, Baptiste Mulot<sup>3</sup>, Florence Levréro<sup>2\*</sup>**

<sup>1</sup> Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000 Rennes, France.

<sup>2</sup> Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES / CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France.

<sup>3</sup> ZooParc de Beauval & Beauval Nature, 41110 Saint Aignan, France

\* Same contribution

Correspondence: L. Pougault, Université Rennes1, Station Biologique de Paimpont, 35380 Paimpont, France.

E-mail address : [loic.pougault@univ-rennes1.fr](mailto:loic.pougault@univ-rennes1.fr) (L. Pougault)

### **Abstract**

Conversely to acoustic structures, call usage in nonhuman primates, and notably the temporal pattern of their call emission, is highly flexible and socially-guided. Different calling patterns may serve different functions and may be used differently by callers according to their social status. Vocal interaction patterns (*i.e.* duet, chorus and turn-taking) have recently attracted a lot of attention and were found to be socially- and temporally-guided and flexibly used in a lot of nonhuman primate species, notably during affiliative encounters. We studied a captive group of chimpanzees and tested the influence of caller characteristics and context of production on their vocal behavior. The vast majority of calls was isolated calls, thus produced outside of vocal interactions. Overlapping calls, involving adult males' *pant-hooting*, represented the core of vocal interactions, while vocal turn-taking was extremely rare. This supports previous findings with wild chimpanzees. Also, the production of series of consecutive calls was predominant and associated with agonism: the higher-ranked the caller, the faster the speed of call delivery and the shorter the mean call duration. The social importance of consecutive

calling was confirmed by the higher rate of audience reaction when the number of calls composing a series increased, suggesting a strategical use of message redundancy. Besides that, social factors weakly influenced most of the other variables (*i.e.* call rates, interaction patterns). This is intriguing comparing with what is typically found in nonhuman primates, including other apes, and questions the potential influence of the chimpanzees' social system that promotes high competition and low cooperation.

**Keywords** Calling patterns; Chimpanzees; Dominance; Vocal interactions; Turn-taking

## Introduction

For a long time, the non-human primates (NHP)'s vocal repertoires were considered as mainly fixed and showing little acoustic variations (Hammerschmidt and Fischer 2008). However, recent studies highlighted that call emission is highly flexible. Flexibility in call usage can be observed when individuals exercise control to call or to remain silent (e.g. Thomas langurs *Presbytis thomasi* Wich and De Vries 2006). This flexibility is also found when in a given context different call types are produced to target different audiences (e.g. chimpanzees Fedurek et al. 2015), when different call types are combined (e.g. Campbell's monkeys *Cercopithecus campbelli* Ouattara et al. 2009; bonobos *Pan paniscus* Schamberg et al. 2016), and when the speed of call delivery within vocal sequences changes with contexts and caller identities (e.g. Campbell's monkeys Lemasson, et al. 2010b)

A caller can produce a unique call or a series of consecutive calls (a given call type is thus repeated several times in a row). The usage of these two calling patterns may depend on contexts and call types but also on the identity and social characteristics of the caller. Isolated calls often occur in the contexts of territory defense, mate attraction, socio-spatial coordination within and between groups, identity and social status advertising (e.g. baboons *Papio cynocephalus ursinus* Fischer et al. 2004; chimpanzees Levréro and Mathevon 2013). Consecutive calls (the

same call type is produced by the same caller calls several times in a row) are rather frequent in contexts conveying intense emotions such as predator detection (e.g. blue monkeys *Cercopithecus mitis* Papworth et al. 2008) or agonistic interactions (e.g. rhesus macaques *Macaca mulatta* Gouzoules et al. 1984ab; chimpanzees Slocombe & Zuberbühler, 2005). Moreover, rather than calling alone, individuals can choose to interact vocally, i.e. at least two callers call in a coordinated way (see Pika et al. 2018). Vocal interactions mainly consist in intra-group affiliative communication (for review see : Grubb et al., 2003; Kappeler & Van Schaik, 2002, but see Da Cunha and Byrne 2006; Schel and Zuberbühler 2012 for example of inter-group vocal interactions), during which interlocutors exert a causal influence on each other (Ravignani et al. 2014).

Three vocal interaction patterns are frequently described in NHP: (i) duetting, (ii) chorusing and (iii) turn-taking. Duetting appears when precisely two individuals synchronize long series of calls or songs with stereotyped temporal association (Langmore, 2002). It generally concerns vocal interactions between two partners of the opposite sex (Bailey 2003) and is common among nonhuman primates with stable monogamy and territoriality (for review see : Fuentes, 1998; Geissmann, 2002). Several elements, such as degree of synchronicity, suggested that the duet song provides information on the strength of the pair bond (Fuentes 1998; Geissmann 2002). Choruses are described as “a simultaneous production of the same call type” by more than two individuals (Yoshida and Okanoya 2005). The function of chorus in primates can be diverse. For instance, the dawn choruses produce by guereza colobus (*Colobus guereza*) appears as a competitive display used by resident males to interact with potential rivals from neighboring groups or to impress females from their own group (Harris et al. 2006; Schel and Zuberbühler 2012). This contrasts with the well described *pant-hoot* choruses in chimpanzee (Mitani and Gros-Louis 1998) occurring between adult males from the same community and used as a short-term bonding signal (chimpanzees Fedurek et al. 2013). Lastly, vocal turn-taking

is defined as the orderly exchange of communicative signals between individuals adjusting the timing of their vocal responses by respecting a minimum silence gap, preventing call overlap, and a maximum silence gap, ensuring call coordination (Pika et al. 2018). Vocal turn-taking was found in a broad range of primate species (Pougnault et al. 2020a). Some authors suggested that cooperation abilities between interlocutors underlie this vocal coordination (Levinson 2016; Takahashi et al. 2013).

The “social bonding hypothesis” suggests that vocal interactions should play a key role in maintaining and reinforcing bonds between individuals in species forming large social groups with tight bonds (Dunbar, 1996). Vocal interactions occur indeed mostly between interlocutors who are non-randomly chosen, depending on, according to species, their affiliative bonds (e.g. humans *Homo sapiens* Cutting & Dunn, 2006; chimpanzees Fedurek et al. 2013, japanese macaques Arlet et al. 2015, bonobos Levréro et al., 2019) and/or respective ages (e.g. humans Bascom, 1942, gorillas Lemasson et al. 2018, Campbell's monkeys Lemasson et al., 2010a). The temporal organization of vocal interactions is also flexible and often socially-guided. The inter-calls delay during vocal turn-taking can be shorter when interlocutors have strong affiliative bonds (squirrel monkeys *Saimiri sciurus* Masataka and Biben 1987). Vocal turn-taking is shaped by comparable social and temporal “rules”, as reported in all human cultures (Ghiglione 1986; Sacks et al. 1974; Stivers et al. 2009).

Researchers have questioned the evolutionary origin and function of these “rules” and suggested the need for more cross-species comparisons (e.g. birds Henry et al. 2015; japanese macaques Sugiura 1993; marmosets *Callithrix jacchus* Takahashi et al. 2013). The growing body of literature about vocal interaction across NHP species suggested that the vocal turn-taking shaped by social and temporal “rules” is strongly represented in primate lineage (Pougnault et al., 2020a). However, the origins of such vocal behaviors in NHP remain debated because of the lack of empirical data, the majority of available studies focusing on monkeys

and neglect great apes. Recently, studies conducted on captive populations of bonobos and gorillas demonstrated that both species engage vocal exchanges respecting a short inter-call delay, preventing call overlap (i.e. temporal organization ; gorillas : Lemasson et al. 2018; Pougnault et al. 2020b; bonobos : Levréro et al. 2019), and occurring between non-random interlocutors (i.e. social organization; partner close in age for gorillas : Lemasson et al. 2018; according to the social affinities in bonobos : Levréro et al. 2019). However, results obtained by Arcadi (2000) several years ago on free-ranging male chimpanzees did not show the existence of such vocal behavior in this great ape species. With respect to the social bonding hypothesis, this finding appears surprising given their high degree of sociality. Chimpanzees are territorial and male-philopatric species living in flexible fission-fusion societies (Nishida 1968). Males are gregarious and form strong social bonds that enable them to deal with high level of inter- and intra-community competition (Duffy 2006; Muller and Mitani 2005), while females tend to live most of the time in smaller family groups on the periphery of the community (Langergraber et al. 2009; Lehmann and Boesch 2009) and sometimes with other females as well in close bonds, despite high variations according to the study sites (Wakefield 2002; Williams et al. 2002). The strong hierarchy among individuals and the high degree of competition in chimpanzees govern most of their social behaviors (deWaal 1982; Mitani 2009; van Lawick-Goodall 1968). Despite their high sociality, they appear to be less prone to cooperative behaviors than their bonobo relatives (Gruber and Clay 2016), for which socially-guided vocal turn-taking was found (Levréro et al. 2019).

Given the social complexity of chimpanzees' communities and based on previous demonstration of vocal turn-taking in great apes (gorillas Lemasson et al. 2018; bonobos Levréro et al., 2019), we propose here to replicate the study conducted by Arcadi (2000) on both male and female chimpanzees housed at the ZooParc de Beauval (n=15), by conducting the same analysis previously applied on gorillas and bonobos (see methods for more details).

First the objective was to confirm or not the potential absence of vocal turn-taking in chimpanzees. If turn-taking occurs, we expected that dominance plays a role in the observed patterns, as in other species where hierarchy is strongly defined like white-faced capuchins (*Cebus capucinus* Digweed et al. 2007). According to previous findings on gorillas and bonobos (Lemasson et al., 2018; Levréro et al., 2019), but also on others non-ape species (e.g. Campbell's monkeys Lemasson et al. 2010a; japanese macaques Sugiura, 1993), vocal turn-taking mainly occurred in peaceful contexts. Thus, the captive condition and the social stability of our study group should be suitable to record such vocal interactions if they matter for the species.

Second our goal was to provide a preliminary description of temporal call usage patterns in this captive group of chimpanzees. As already shown by previous studies (e.g. rhesus macaques *Macaca mulatta* Gouzoules et al. 1984ab; Campbell's monkeys Lemasson et al. 2010b; blue monkeys Papworth et al. 2008), we expected to find a high proportion of consecutive calls in contexts perceived as associated with high arousal such as aggressive interactions (rhesus macaques Gouzoules et al., 1984ab), and an influence of hierarchical rank on the temporal features (i.e. duration of calls and inter-calls delay) of the consecutive calls.

Third we tested the impact of the different temporal calling patterns on the non-vocal reaction of the audience in order to question the possible social function of the different calling patterns found. We expected to observe stronger reactions following insistent consecutive calling (rhesus macaques Gouzoules et al., 1984ab; japanese macaques Koda, 2004).

## Methods

### *Compliance with ethical standards*

All applicable international, national and institutional guidelines for the care and use of animals were followed. The study has been conducted in accordance with the current laws in France



(agreement with the 2010/63/UE). Animal husbandry and care were under the management of the animal caregivers of the ZooParc de Beauval, France. Our study was based on non-invasive observations and was approved by the ethics committee “Comité Rennais d’Ethique en matière d’Expérimentation Animale” (*i.e.* Rennes Ethical Comity for Experiments using Animals; CNREEA 2018041710224608-APAFIS#14920).

### ***Animals and housing conditions***

The captive studied group was formed in 1993. In 2018, it included fifteen chimpanzees. We defined the age classes following Reynolds (2005) and we considered the sub-adult, juvenile and infant individuals as the “immature class” (14 years of age). According to this classification, the group composition was nine adults (two males and seven females), and six immature individuals (four males and two females) (see Table 1). All adult females were unrelated and paternity could not be confirmed but all immature individuals were sired by one of the two current adult males.

The animals were housed at the ZooParc de Beauval (France) in an enriched indoor enclosure of 150m<sup>2</sup> (5.30m high, with glass walls, wired roof, straw litter and wood perches, as well as temperature and humidity control), connected to several adjacent cages totaling 20m<sup>2</sup> (average size: 5m<sup>2</sup>). Depending on the weather conditions, group had also access to a 1200m<sup>2</sup> outdoor enclosure (with grass, bush and wood perches).

### ***Data Collection***

Social and vocal data were collected from June to November 2018 (excluding nine weeks during summer to avoid the season with the highest visitor presence rate). All the data were collected by one observer (LP) from either a footbridge (social and vocal data) above the indoor enclosure (Y-shaped, length 6.5m for each arm on which the observer could walk) or next to

the outdoor enclosure (social data only) according to the location of the group. Chimpanzees had been habituated to the presence of the observer and recording equipment for 15 days before the data collection. Behavioral and vocal recordings were conducted between 9 am and 6 pm.

#### *Social data collection*

The social data was obtained using two sampling methods: scan sampling (every 10 minutes) to score inter-individual distances, and individual focal sampling (for 10 min) to score social interactions (Altmann 1974). In total 108 scans were performed where all the group members were visible and were used for the following analysis. Each pair of chimpanzees was scored as “In” (i.e. within hand reach, inspired from Funkhouser et al., 2015) or “Not In” spatial proximity. Concerning focal sampling, each group member (n=15) was observed 1h30 (in random order). We classified the social interactions into three categories (Pollick & De Waal, 2007): aggression (displace, charge, wrestle, hit, grab, bite, steal food and supplant), submission (avoid calmly and flee), affiliation (groom, play with physical contacts, touch, smell and embrace).

#### *Social data processing*

We calculated three indices at a dyadic level (for each pair of individuals AB): (i) the dominance index, (ii) the level of social integration and (iii) the level of spatial proximity.

### **Dyadic indices**

#### **Dominance index**

Following a calculation proposed by Zumpe and Michael (1986) and Lemasson and Hausberger (2004), a dominance index (in percentage) was calculated for each group member on the basis of agonistic interactions, excepted for two infants as they were still dependent on their mothers. For all the recorded interactions, the aggressor and the receiver could be clearly identified (N=

115). For each pair of individuals, (i) the number of aggressive behaviors displayed by one individual to the other one was expressed as a percentage of the total number of their mutual aggressive behaviors, (ii) the number of displayed submission behaviors was calculated with the same method, and (iii) the two previous scores were averaged for each individual.

#### Level of social integration

To quantify the level of social integration, we used the Dyadic Affiliation Index (D.A.I.) previously developed by Pollick and De Waal (2007), and adapted by Prieur and colleagues (2016). This index allowed assessing the relationship quality based on the relative frequency of affiliative and agonistic behaviors within dyads. This index increases with affinity, starting from 0 in the absence of affinity. It is calculated as:

$$D.A.I._{AB} = \frac{\left( \sum_{i=1}^n \frac{f_{iAB}}{f_i} \right)}{1 + \frac{\left( \sum_{j=1}^{n'} \frac{h_{jAB}}{h_j} \right)}{n'}}$$

where  $f_{iAB}$  is the total number of affiliative interactions of the behavior ( $i$ ) expressed by  $A$  towards  $B$ ;  $f_i$  is the mean number of affiliative interactions of the behavior ( $i$ ) across all dyads;  $n$  is the number of affiliative behaviors expressed by  $A$  towards  $B$ ;  $h_{jAB}$  is the total number of agonistic interactions of the behavior ( $j$ ) expressed by  $A$  towards  $B$ ;  $h_j$  is the mean number of agonistic interactions of the behavior ( $j$ ) across all dyads;  $n'$  is the number of agonistic behaviors expressed by  $A$  towards  $B$ .

#### Level of spatial proximity

This index corresponded to the number of scan samples in which  $A$  and  $B$  were found within hand reach to each other (namely in spatial proximity) out of the total number of scans where both  $A$  and  $B$  were visible (e.g. Lemasson et al. 2018).

**Individual indices**

Lastly, five social indices were considered at individual level (see Table 1). Three of them were obtained from the social indices previously calculated at the dyadic level: (i) the dominance index (*i.e.* average of all dominance dyadic indices involving a given individual; Lemasson et al., 2018), (ii) the level of social integration (*i.e.* average of all D.A.I. involving a given individual), (iii) the level of spatial proximity (*i.e.* average of all spatial proximity dyadic indices involving a given individual). In addition, the age and sex of each individual were taken account.

**Table 1** | Composition of the studied group and social position of individuals in the group

Subject name (birth date)	Age-sex category	Mother's name	Dominance index	Level of social integration	Level of spatial proximity
Joseph (1974 or 1975 <sup>†</sup> )	Adult male	-	3	0.12	96
Charlotte (01/08/1976)	Adult female	-	0	0.78	199
Micheline (1977 or 1978 <sup>†</sup> )	Adult female	-	1.5	1.80	100
Baraka (02/01/1979)	Adult female	-	1.75	4.38	126
Julie (1981 or 1982 <sup>†</sup> )	Adult female	-	0.5	0.33	126
Bonobo (1981 or 1982 <sup>†</sup> )	Adult female	-	1.5	0.36	148
Gypso (25/02/1987)	Adult female	-	1	0.36	96
Gamin (1988 or 1989 <sup>†</sup> )	Adult male	-	0	1.02	63
Domi (04/12/1989)	Adult female	-	2	2.04	239
Sangha (21/07/2006)	Subadult female	Bonobo	0.5	0.79	113
Wamba (23/11/2008)	Juvenile female	Bonobo	0.5	1.74	98
Tumba (14/01/2009)	Juvenile male	Domi	0.5	2.92	129
Lukombe (03/08/2011)	Juvenile male	Bonobo	0.75	3.04	145
Yumbi (05/02/2014)	Infant male	Bonobo	-	12.88	136
Lobaï (11/06/2014)	Infant male	Domi	-	12.93	218

<sup>†</sup> Uncertainty regarding the date of birth of these individuals, estimation  $\pm 1$  year.

*Vocal data collection*

Calls were recorded using a Marantz PMD660 handheld digital audio recorder (Marantz, Japan) (sample rate 44.1kHz, resolution 32bits, wav. format) connected to a directional Sennheiser MKH70-1 microphone (Sennheiser, Germany), and a directional AV-JEFE TCM 160 LAVALIER tie microphone (AV-JEFE, Singapore) for the experimenter comments. We recorded vocalizations of all group members on an all-occurrence basis during recording sessions lasting 1h30 each per day (Altmann 1974). In total, 592 call utterances were collected during the 26 hrs of vocal recordings. Due to uncertainty in caller identity and / or to poor recording quality, 13% of these call utterances were discarded.

To determine the call types, we referred to previously published vocal repertoire done by Slocombe and Zuberbühler (2010) and added the “*Hoo*” vocalization recorded during our data collection and previously described by Goodall (1986) (see Table SI in Electronic Supplementary Materials for a complete list of call types that we used). Also, we defined a “*mixed-call*” when different call types were mixed together into the same call utterance (Lemasson et al., 2018). Nevertheless, due to a too low sample size, we chose to conduct our investigation of temporal calling patterns at the vocal repertoire level, regardless of the type of calls involved. Further studies will be needed to go deeper into the patterns that are observable at the call type level. As an indication, the percentage of each call type produced is given in the Table SI in Electronic Supplementary Materials.

*Vocal data processing*

We investigated the temporal organization of the call production within the group measuring all the inter-call durations thanks to the spectrograms generated by PRAAT software version 6.1.09 using a Gaussian window.

For processing about turn-taking, we consider that the inter-call duration corresponded to the silent gap between two consecutive calls of any call types produced by two different identified callers, from the offset of the first call to the onset of the second call (as in Lemasson et al. 2018 and Levréro et al. 2019). To define if a call belonged to a vocal interaction (*i.e.* which corresponded to at least two successive calls of alternating individuals with a short or a negative inter-call delay) or not, we used a graphical method already applied to define the maximum response delay for a vocal interaction in several primate species (e.g. Lemasson, et al., 2010; Lemasson et al., 2018; Levréro et al., 2019). We defined a temporal window of a maximum of 20 seconds after the onset of the first call, which was conservative regarding other findings in multiple species (e.g. Koda, 2004; Lemasson, Gandon, et al., 2010; Lemasson et al., 2018; Levréro et al., 2019; Sugiura, 1998) to highlight an eventual pattern of vocal interaction. We then represented the distribution of inter-call delays and defined the temporal threshold to classify whether a call belonged to a vocal interaction or not. The threshold indicated the maximum response delay for vocal interaction. It was placed at the location where inter-call delays started rising above a baseline representing a background noise of inter-call delay out of vocal interactions. When several calls overlapped (negative inter-call durations), we conserved only the calls where both callers were identifiable.

We were thus able to quantify the proportions of: (i) call utterances produced during vocal interactions, and (ii) call utterances produced in absence of vocal interaction, which consisted of calls produced independently from the vocal utterances of other group members (see also Bouchet et al. 2010; Levréro et al. 2019).

Call utterances from one individual can also be produced following two different temporal calling organizations, irrespective of whether or not they belonged to a vocal interaction: a single call or the same call type repeated and separated by short silent gaps, thus composing a series of calls. The same temporal threshold as the one defined for vocal interactions was

applied to classify whether a call was produced alone or in a series (Koda, 2004; Lemasson et al., 2013, 2018). For clarity, the two different temporal calling organizations will be respectively named: (i) “Unique Call” (hereafter UC), (ii) “Series of Calls” (hereafter SC). It is worth noting that in the unique cases of the pant-hoot and pant-grunt vocalizations, UC could be composed of several unmerged sound units because pant-hoots and pant-grunts are, by definition, multi-unit calls incorporating noisy inhalation phases (Gros-Louis & Mitani, 1998; Laporte & Zuberbühler, 2010). Following the Koda (2004)’s suggestion saying that consecutive calls elicited more vocal response from the audience than unique call, we investigated the proportions of UC and SC in the vocal interaction pattern and we searched whether UC and SC were more in an “initiator” or “respondent” position in the interactions.

Lastly, we studied the production of calling patterns based on the hourly rates of calling according to the individual characteristics above-mentioned (*i.e.* age-sex classes, dominance index, level of social integration and level of spatial proximity). We noted the context in which the emitters produced their calls in terms of activities and behaviors (see Table 2). During the 10 seconds following each vocal production, we recorded the non-vocal reaction of the audience within 10 to 15 meters from the emitter (neutral social behaviors: looking towards the emitter or the other group members; walking in the direction of the emitter; positive social behaviors: embrace, grooming, sharing food; negative social behaviors: fleeing, defense, refusal of physical contact or attacking the emitter).

**Table 2** | Definition of the behavioral contexts recorded during vocal communication in the captive chimpanzee group in the ZooParc de Beauval

Context	Definition
Affiliative	The emitter is involved in a social positive physical contact with another. Contact-play behaviors are included as affiliative context.
Aggressive	The emitter chases away or physically attacks another one.
Submissive	The emitter avoids calmly another individual or runs away.
Frustration	The emitter is facing another individual who refuses to share food or to interact sexually.
Social meeting	The emitter is joining one individual or a group.
Feeding	The emitter is foraging, manipulating food, eating or in position to drink.
Waiting for food	The emitter saw/heard the caregivers into the corridors with food just before food distribution.
Resting	The emitter remains stationary without participating in any self-directed (including self-grooming) or other-directed physical activity; includes sleeping.

### **Data Analysis**

Normality and homogeneity of variances were assessed by inspecting residuals following Shapiro-Wilk W tests. Given the non-normality of data, we ran non-parametric tests (Siegel and Castellan 1988) with a significance threshold of 0.05. Data description was done using percentages, means and standard deviations.

The individual hourly rates of call production were used to investigate the potential influence of social factors on calling activity: sex (Mann-Whitney test), age, dominance index, level of social integration and spatial proximity (Spearman correlations). The individual temporal call production (i.e. UC and SC) was described in the same way and Wilcoxon signed rank-tests were applied to determine the most likely contexts of emission of both patterns. Binomial tests allowed us to tests the favored position of UC and SC (“initiator position” vs “respondent position”) during vocal interactions. A generalized linear mixed model (GLMM) with a negative binomial distribution was conducted to define the contexts in which the SC had the most average number of calls (dependent variable: average number of calls in SC; fixed effect: context of production; and random factor: individual caller identity; *lme4* package in R).



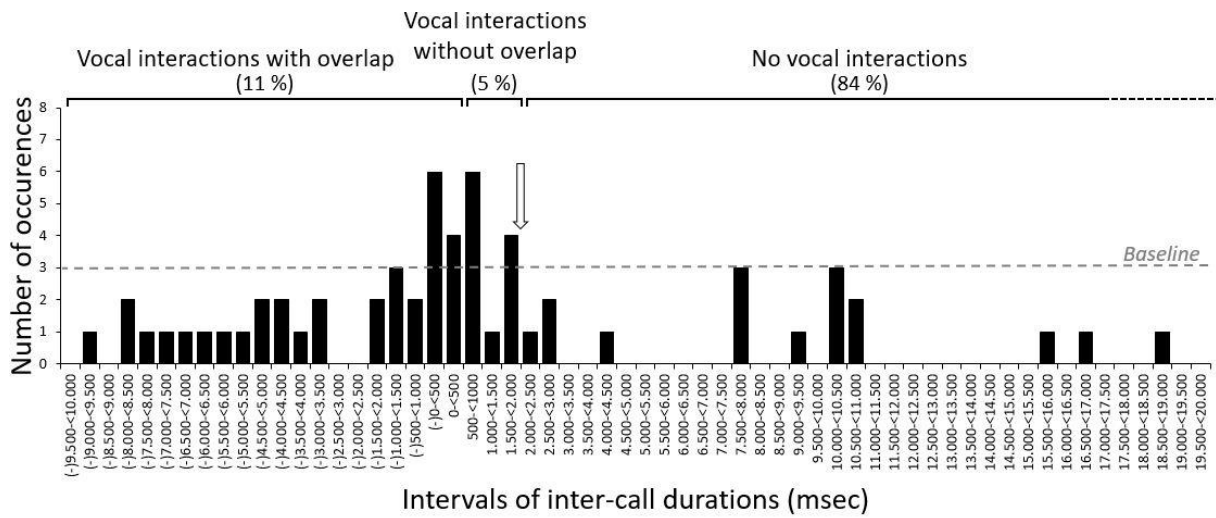
Finally, the influence of the number of calls produced in SC on the non-vocal reaction of the audience was investigated using a Spearman correlation, and the potential influence of social factors on the temporal structure of SC (i.e. average delay between two consecutive calls, as well as the average duration of calls) was evaluated : sex (Mann-Whitney test), age, dominance index, level of social integration and spatial proximity (Spearman correlations).

## Results

### 1. Production of calling patterns

#### 1.1. Temporal features of calling patterns

From all the recorded calls for which the caller was identified (N=516), 120 occurred within a time window of 20 secs of another call (see *Methods*); the remaining 396 calls occurred more than 20 secs from another call. From these 120 calls, we studied the distribution of inter-call durations (n= 60 inter-calls duration). We graphically defined a positive inter-call delay of a maximum of 2 s between two successive calls produced by two different callers (Fig. 1) to define a vocal interaction. Out of the 516 calls recorded, only 16% were produced as part of a vocal interaction. Among the vocal interactions, we observed some negative inter-call durations that corresponded to call overlaps. Overlapped calls occurred in 63% (52/82) of the calls produced in vocal interactions. They corresponded to 14 vocal interactions and were produced on average by  $3.71 \pm 0.83$  callers. Vocal interactions with overlap were on average longer ( $26.22 \pm 34.61$  s) than vocal interactions without ( $7.82 \pm 7.57$  s). Vocal interactions without overlap represented 37% (30/82) of vocal interactions, corresponding to 15 vocal interactions, and were produced on average by  $2.0 \pm 0.0$  callers. Our data did not allow knowing if all the recorded vocal interactions could be considered as purely communicative, *i.e.* whether individuals exert a causal influence on each other or not.



**Figure 1** | Distribution of inter-call durations between two successive calls from different callers recorded from all occurrence samples ( $n = 60$  inter-call durations recorded in a time window of 20 secs after hearing a call). Grey dotted line corresponds to the baseline of the occurrence of positive inter-call durations. White arrow corresponds to the threshold indicating the maximum response delay for vocal interaction, and was placed at the location where inter-call delays start rising above the baseline when moving from long delays to the peak observed at around two seconds.

All call types did not contribute in the same way to the different calling patterns (Table 3). Indeed, *pant-hoots* were the most frequent call types produced during the vocal interactions (with overlap: 72%, 37/52; without overlap: 37%, 11/30) while *whimpers* and *screams* were the most produced (respectively 31% 134/434, and 26% 112/434) in the context of no vocal interactions.

**Table 3** | Contribution of different call types in the three calling patterns

Call types	Calls in vocal interactions		Calls out of vocal interactions (n= 434)
	With overlap (n= 52)	Without overlap (n= 30)	
<b>Bark</b> (n= 7)	0 %	7 %	1 %
<b>Cough</b> (n= 43)	0 %	17 %	7 %
<b>Grunt</b> (n=43)	0 %	13 %	10 %
<b>Hoo</b> (n=25)	0 %	10 %	5 %
<b>Laugh</b> (n=36)	0 %	0 %	9 %
<b>Mixed-Call</b> <sup>†</sup> (n= 6)	5 %	3 %	1 %
<b>Pant-Grunt</b> (n= 7)	9 %	3 %	1 %
<b>Pant-Hoot</b> (n= 60)	72 %	37 %	6 %
<b>Scream</b> (n= 131)	14 %	3 %	26 %
<b>Squeak</b> (n= 21)	0 %	0 %	5 %
<b>Whimper</b> (n= 137)	0 %	10 %	31 %

<sup>†</sup>Mixed calls are combinations of different call types.

## 1.2. Call production rate according to the characteristics of individuals

### 1.2.1. Influence of age-sex classes

Regarding calling patterns, all individuals called more in the context of no vocal interactions than of vocal interactions (Table 4). The two adult males contributed to some extent to all the calling patterns, while the other age-sex classes almost never interacted vocally.

**Table 4** | Hourly rate of vocal production of the different age-sex classes to the three calling patterns.

Age-sex category	Vocal interactions		No vocal interactions (mean ± S.D.)
	With overlap (mean ± S.D.)	Without overlap (mean ± S.D.)	
<b>Adult males</b> (n= 2)	0.25 ± 0.14	0.06 ± 0.03	0.73 ± 0.65
<b>Adult females</b> (n=7)	0.08 ± 0.06	0.03 ± 0.02	0.59 ± 0.96
<b>Immature males</b> (n=4)	0.07 ± 0.04	0.07 ± 0.07	2.26 ± 2.76
<b>Immature females</b> (n= 2)	0 ± 0	0.02 ± 0.03	0.90 ± 0.95

### 1.2.2. Influence of social indices

The dominance index, the level of social integration and spatial proximity had no influence on the individual hourly vocal production rate whatever the calling patterns (vocal interactions

with overlap:  $\rho_{\text{dominance index}} = 0.16$ ,  $P = 0.60$ ;  $\rho_{\text{social integration}} = -0.44$ ,  $P = 0.10$ ;  $\rho_{\text{spatial proximity}} = 0$ ,  $P = 1$  respectively; vocal interactions without overlap :  $\rho_{\text{dominance index}} = -0.082$ ,  $P = 0.77$ ;  $\rho_{\text{social integration}} = -0.0924$ ,  $P = 0.74$ ;  $\rho_{\text{spatial proximity}} = 0.057$ ,  $P = 0.84$ ; context of no vocal interactions :  $\rho_{\text{dominance index}} = -0.20$ ,  $P = 0.7$ ;  $\rho_{\text{social integration}} = -0.123$ ,  $P = 0.66$ ;  $\rho_{\text{spatial proximity}} = -0.13$ ,  $P = 0.64$ ).

## 2. Individual vocal production: Unique Call or Series of Calls

### 2.1. Contribution of the Unique Calls and Series of Calls to the calling patterns

Individuals uttered either a Unique Call (UC), which represented 21% (108/516) of the total call production, or a Series of Calls (SC, 79%, 408/516), each call separated by a short silent gap (<2 secs) and represented 57 series. The number of calls produced consecutively (forming a SC) by a single individual ranged from two to 38 calls (mean  $7.19 \pm 7.52$ ).

UC were produced in same proportions during vocal interactions than in the absence of vocal interactions (Binomial test,  $N_{\text{vocal interactions}}: 49$ ,  $N_{\text{absence of vocal interactions}}: 59$ ,  $P=0.39$ ), whereas SC were significantly more produced in the absence of vocal interactions ( $N_{\text{vocal interactions}}: 7$ ,  $N_{\text{absence of vocal interactions}}: 50$ ,  $P<0.001$ ). UC accounted for 82 % of vocal interactions with overlap ( $N_{\text{UC}}: 28$ ,  $N_{\text{SC}}: 6$ ,  $P<0.001$ ) and 84% of interactions without overlap ( $N_{\text{UC}}: 21$ ,  $N_{\text{SC}}: 4$ ,  $P<0.001$ ).

### 2.2. Contribution of the call types to UC and SC

UC were mostly composed of *pant hoots* (52%, 56/108 Table 5) whereas more diverse call types composed SC, such as *whimpers* (30%, 17/57), *screams* (23%, 13/57), *cough* (18%, 10/57), *Hoo* (12%, 7/57) and *grunts* (11%, 6/57; see Table 5). Some call types such as *barks*, *pant-hoots*, *pant-grunts* and *mixed calls* were only produced in UC, while *laughs* and *squeaks*, which were rare calls, were exclusively recorded into SC (Table 5).

**Table 5** | Contribution of the different call types produced alone ('Unique Calls') and in 'Series of Calls' (same call type repeated and separated by short silent gaps of a maximum of 2 secs).

Call types	Unique Call (n= 108)	Series of Calls (n= 57)
<b>Bark</b>	6 %	0 %
<b>Cough</b>	7 %	18 %
<b>Grunt</b>	8 %	11 %
<b>Hoo</b>	9 %	12 %
<b>Laugh</b>	0 %	2 %
<b>Mixed-Call*</b>	5 %	0 %
<b>Pant-Grunt</b>	6 %	0 %
<b>Pant-Hoot</b>	52 %	0 %
<b>Scream</b>	5 %	23 %
<b>Squeak</b>	0 %	5 %
<b>Whimper</b>	2 %	30 %

### 2.3. Influence of the individuals' characteristics in UC and SC production rate

#### 2.3.1. Influence of age-sex classes

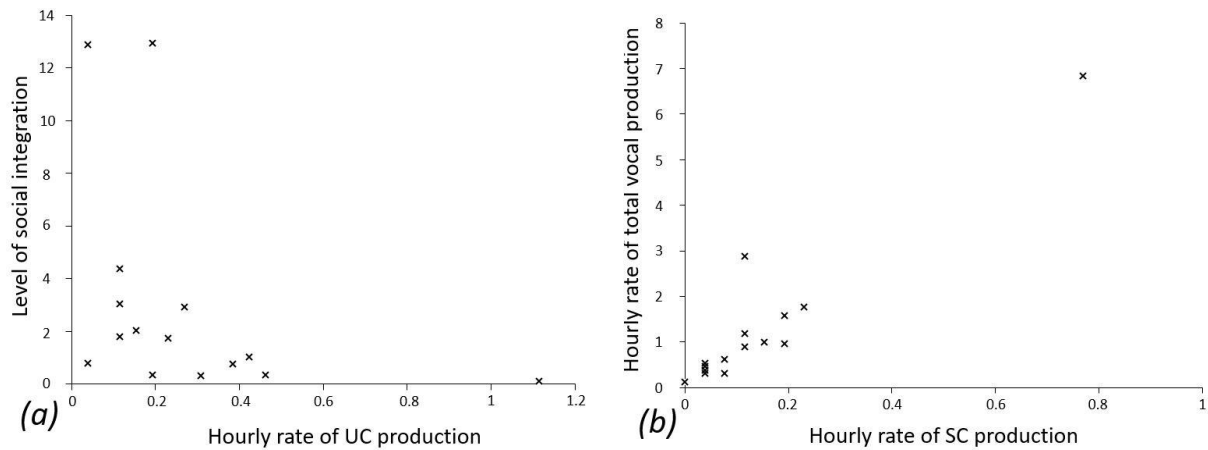
Among all the age-sex classes, adult males had the highest hourly production rate of UC ( $0.76 \pm 0.49$ , N= 2), which were three times more than adult females ( $0.24 \pm 0.14$ , N=7). Immature individuals, both males (N=4) and females (N=2), produced the lowest hourly rates of UC ( $0.15 \pm 0.1$  and  $0.13 \pm 0.14$  respectively). Different results were observed for the SC: immature males had the highest hourly rate ( $0.31 \pm 0.31$ ). There was no difference between adult males and immature females ( $0.13 \pm 0.14$  and  $0.13 \pm 0.08$  respectively) and adult females produced the least of SC ( $0.06 \pm 0.04$ ).

#### 2.3.2. Influence of social indices

We found that the higher the level of social integration of individuals, the lower the hourly production rate of UC (Spearman correlation,  $\rho = -0.62$ ,  $P < 0.05$ , Fig 2a).

No other social indices influenced the production rate of UC ( $\rho_{\text{dominance index}} = -0.076$ ,  $P = 0.81$ ;  $\rho_{\text{spatial proximity}} = -0.38$ ,  $P = 0.16$ ). Similarly, no social indices impacted the production of SC ( $\rho_{\text{dominance index}} = -0.017$ ,  $P = 0.57$ ;  $\rho_{\text{social integration}} = 0.091$ ,  $P = 0.79$ ;  $\rho_{\text{spatial proximity}} = 0.081$ ,  $P =$

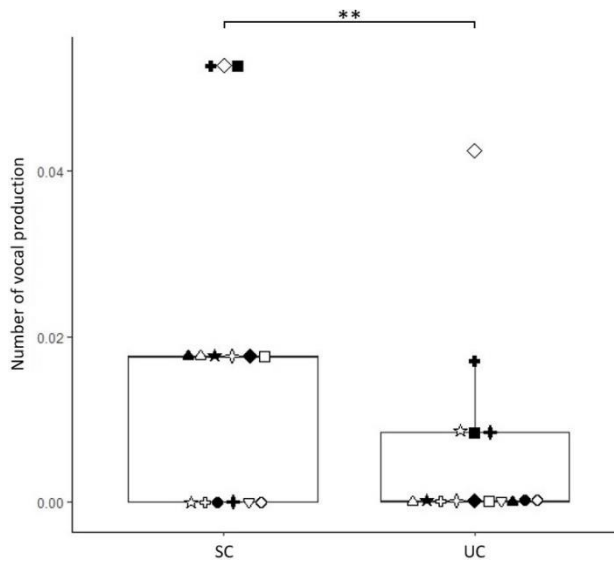
0.77). Individuals with the highest hourly vocal production rate were however the individuals with the highest hourly production rate of SC ( $\rho=0.87$ ,  $P<0.001$ , Fig 2b), while no such correlation was found with the production of UC ( $\rho=0.40$ ,  $P=0.14$ ).



**Figure 2** | (a) Hourly rate of UC production negatively correlated with the level of social integration of individuals (Spearman correlation,  $P<0.05$ ). (b) Hourly rate of SC production positively correlated with the hourly rate of total vocal production of individuals (Spearman correlation,  $P<0.001$ ).

#### 2.4. Influence of the contexts on the production of UC and SC

In the “aggressive” context, SC were more frequent than UC (Wilcoxon signed rank test,  $V=63$ ,  $P<0.01$ , Fig 3). For all the other contexts, SC and UC were equivalently produced (Wilcoxon signed rank test,  $V_{\text{affiliative}}=19$ ,  $P=0.08$ ;  $V_{\text{submissive}}=12$ ,  $P=0.27$ ;  $V_{\text{frustration}}=0$ ,  $P=1$ ;  $V_{\text{feeding}}=5$ ,  $P=0.42$ ;  $V_{\text{food waiting}}=7$ ,  $P=0.066$ ;  $V_{\text{social meeting}}=5$ ,  $P=0.28$ ;  $V_{\text{resting}}=14$ ,  $P=0.054$ ).



**Figure 3** | Distribution of the series of calls (SC) and unique calls (UC) in “aggressive” context. The different individuals are represented by a different symbol ( $N=15$ ). Box and whisker plots report the median, 25th and 75th percentiles, and the lowest and highest data, which were no more than 1.5 interquartile range from the boxes (Wilcoxon signed rank test, \*\*:  $P < 0.01$ ).

Additionally, the average number of calls composing the SC was higher in the “aggressive” and “resting” contexts than in other contexts (GLMM, for comparisons including the “aggressive” or “resting” contexts:  $Z$  ranged from 2.142 to 2.703,  $P < 0.05$ ; for all others:  $Z$  ranged from 0.026 to 1.675, for all  $P > 0.05$ ; Fig SI in Electronic Supplementary Materials).

### 2.5. Influence of the individuals’ characteristics on the temporal features of SC

In “aggressive” context more dominant individuals presented shorter call intervals in SC (Spearman correlation,  $\rho = -0.80$ ,  $P < 0.01$ , Fig SIIIa in Electronic Supplementary Materials), and also a lower mean call duration ( $\rho = -0.66$ ,  $P < 0.05$ , Fig SIIIb in Electronic Supplementary Materials).

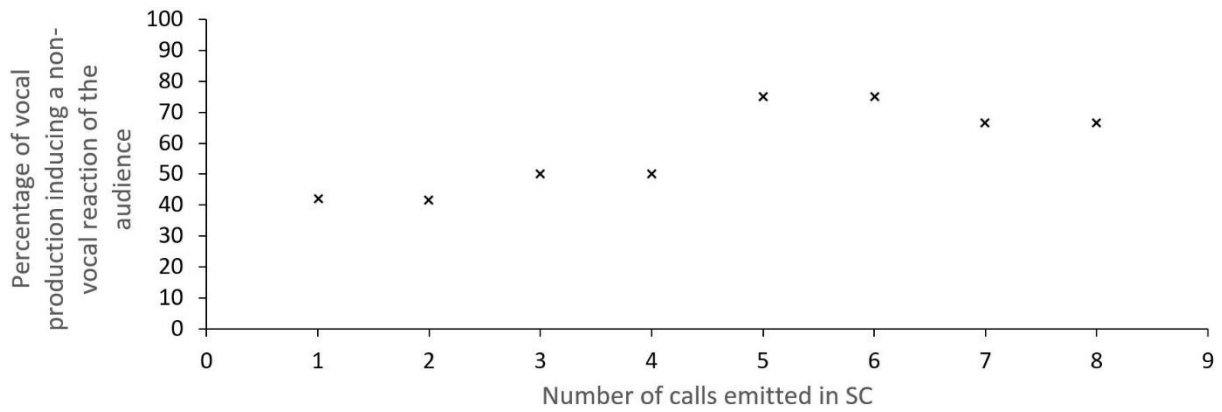
No other correlation between individuals' characteristics and SC features were found neither in the "aggressive" context (Spearman correlation, level of social integration :  $\rho_{\text{delay}} = -0.34$ ,  $P = 0.2186$ ;  $\rho_{\text{duration}} = -0.45$ ,  $P = 0.094$ ; level of spatial proximity :  $\rho_{\text{delay}} = -0.22$ ,  $P = 0.42$ ;  $\rho_{\text{duration}} = -0.29$ ,  $P = 0.3$ ; Age :  $\rho_{\text{delay}} = 0.089$ ,  $P = 0.75$ ;  $\rho_{\text{duration}} = 0.17$ ,  $P = 0.54$ ; Wilcoxon rank sum test, Sex:  $W_{\text{delay}} = 21$ ,  $P = 0.50$ ;  $W_{\text{duration}} = 26$ ,  $P = 0.95$ ), nor in the seven other contexts (*i.e.* affiliative, submissive, frustration, feeding, food waiting and resting; see Table SII in Electronic Supplementary Materials).

### *3. Factors influencing the reaction of the audience*

If we considered all the vocal productions ( $N_{\text{UC}} = 108$  and  $N_{\text{SC}} = 57$ ), we noticed that 36% (60/165) of them induced a non-vocal reaction of the audience. Regarding the remaining 64% (105/165), 8% (13/165) were only vocal reaction and 52% (85/157) did not induce any overt reaction from the visible audience. The main observed non-vocal behaviors of the audience were fleeing (32%, 19/60), looking at the caller (19%, 11/60), attacking the caller (14%, 8/60) or looking at other group members (6%, 4/60). Other non-vocal reactions of the audience were low represented (5%, 3/60 for: swagger, contact-play and refusal a physic interaction; 2% 1/60 for: embrace, grooming, approach to the emitter, defense and sharing food).

We focused here on SC composed of less than nine consecutive calls, given the low occurrence beyond this threshold, which represented 82 % (47/57) of total SCs and totalized 181 consecutive calls. We found that the audience significantly reacted more often when the number of calls in SC increased (Spearman,  $\rho = 0.7637$ ;  $P < 0.05$ , Fig 4).





**Figure 4** | Percentage of vocal production inducing a non-vocal reaction of the audience related to the number of consecutive calls (Spearman correlation,  $P < 0.05$ ).

## DISCUSSION

The calls produced in the studied captive group did not typically elicit vocal responses from other group members. Vocal interactions mostly involved males' overlapping *pant-hoots*. Calls uttered in a series concerned mainly immature males and in aggressive contexts. Also, the temporal features of consecutive calling (*i.e.* duration of calls and call intervals) were negatively correlated with the social status of the caller (*i.e.* dominance index). Moreover, the more calls repeated, the more chances a caller would receive a non-vocal response from a group member. Lastly, contrary to most of the other studied social primates, we found that the vocal production rate and the temporal vocal patterns were weakly influenced by social factors, notably affiliative bonding.

Because chimpanzees are highly social, and following Dunbar (1996)'s hypothesis, we expected a much higher rate of vocal interactions (*i.e.* duet, chorus or turn-taking) in that species. The chimpanzees live in communities with 20 to 120 members. However, their society are well known to be driven by a high degree of competition (*e.g.* Muller and Mitani 2005; Nishida 1968) and they appear less cooperative than bonobos (*e.g.* Gruber and Clay 2016) who do use higher rates of temporally and socially organized vocal interactions (Levréro et al. 2019).

Our results show that the frequency of the emission of temporally and socially organized vocal interactions seems not exclusively connected to the group size and the presence of tight bonds between group members, but may also be influenced by the nature of the bonding network, which is similar for bonobos and chimpanzees, and strength of the dominance hierarchy, dissimilar between these two apes.

Vocal interactions of the studied animals consisted mainly in *pant-hoot* overlap. In the wild, *pant-hoot* chorusing is associated with a clear communication function, notably to form and strengthen alliances between adult males (Fedurek et al. 2013; Gros-Louis and Mitani 1998) (Mitani & Gros-Louis, 1998). Moreover, rivals seem to be able to perceive the number of participants to a *pant-hoot* chorus and to adapt their behavior in consequence (Wilson et al. 2001). On the contrary, vocal interactions based on turn-taking (*i.e.* without overlap) were rare in our study. This supports previous findings in chimpanzee Kibale population (Arcadi, 2000), despite the fact that our study investigated all the possible call types and callers and not only males. Arcadi (2000) collected  $9.0 \pm 5.5\%$  call “responses” (5% in our case) with a response delay arbitrarily defined at 5 s maximum (graphically defined at 2 s in our case). This makes chimpanzees a very atypical primate species in this regard. Vocal turn-taking with overlap avoidance is found in all human cultures (Stivers et al. 2009), even in a few days old baby (Stern et al. 1975), and is very common in New World and Old World monkeys and apes (Pougnault et al., 2020a). A part of the reason of their particular vocal behavior may be due to their social system and the high degree of competition. We indeed know that chimpanzees do take turn when it comes to visual communication (Fröhlich et al. 2016), so this is definitely not a cognitive issue. Our analysis was strictly vocal and we acknowledge that it is difficult to understand the interactions properly in chimpanzees without considering their gestural communication (Graham et al. 2018; Roberts and Roberts 2019; van Lawick-Goodall 1968). The fact that we did not find any influence of individual and social characteristics of callers on

the rate of interactions may result from the composition of the studied group (not all age-sex classes were equally represented) and the low diversity of social contexts encountered by individuals. More comparative work is needed to explore further this question.

Nonetheless, in the “aggressive” context, aggressors favored the production of series of calls (rather than unique calls). Because these behaviors were often associated with harassment from immature males, we can hypothesize that this is a way to vocally express emotional changes. The coding of emotions (notably the negative ones) by the number of vocal utterances per unit of time is a well-known phenomenon in human (Scherer 2003) and nonhuman primate species (Owren and Rendall 1997). In chimpanzees, this behavior was also previously described in cases of high arousal such as discovery of good-quality food (Dittus, 1984; Hauser & Marler, 1993). The possibility that this behavior is however also used as a social strategy cannot be excluded. So, the use of consecutive calls in high arousal contexts appeared to have the ultimate purpose of attracting the attention of the other group members.

In line with the previous finding, we found that the social status of the emitter during aggression influenced the temporal features of the series of calls. The higher-ranked the caller, the faster the speed of call delivery and the shorter the mean call duration. It has already been found that in species with strong hierarchy, such as baboons, the social status of the caller can be encoded vocally, either in the call’s acoustic structure (Fischer et al. 2004), in the overall call rate (chimpanzees Mitani & Nishida, 1993), or in the preferential use of some call types (baboons Cheney et al. 1995). The social importance of consecutive calling was also confirmed by the fact that the audience reacted more often when the number of calls composing a series increased, highlighting a possible strategical use of message redundancy. The use of consecutive calls to catch group members’ attention has been discussed in several species (*e.g.* meerkats *Suricatta suricatta* Manser, 2001; meerkats Manser et al., 2002; japanese macaques Koda 2004; humans Golinkoff, 1986). We again acknowledge that the low importance of social

factors on call rates and interaction patterns could be due to our group composition, its stability, and the rarity of some contexts in captivity.

In conclusion, it is admitted that the ability to interact vocally with other could have been an important step in the evolution of vocal communication in primates (Dunbar, 1996). This capacity seems even more crucial in mediating social bonds in a highly flexible society, as in chimpanzees or humans. A low-cost bonding behavior such as coordinated vocal display might be adaptive in such fluid societies (Dunbar, 1996). The function, as well as the content of information in vocal interactions remains unclear and still needs to be explored more systematically. Additionally, the influence of the group composition and the living environment must not be underestimated and can drive the use of vocal patterns. However, our study in captivity, coupled with Arcadi's (2000) similar findings in the wild, draw a picture for chimpanzees markedly different from the orderly vocal behavior of bonobos (Levrero et al 2019), the other closest living parent of humans. All together, these studies provide original insights to think about the social pressure and needs that could lead to vocal turn taking in the framework of human communication evolution research.

## Acknowledgments

We are deeply grateful to the ZooParc de Beauval and the chimpanzee caregivers for their valuable assistance. We also warmly thank Katie Collier for correcting our English. This work was financially supported by a CIFRE grant (#2018\_00074) funded by the Association Beauval Nature pour la Conservation et la Recherche, Centre National de la Recherche Scientifique, Rennes 1 University and Lyon/Saint-Etienne University.

## Authors' contributions

LP, AL, BM and FL conceived, designed and coordinated the study. LP collected data and carried out the statistical analyses. LP, AL and FL interpreted and drafted the manuscript. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

## Competing interest

The authors declare no competing interests

## Data availability

The data generated and analysed during the current study are available from the corresponding author on reasonable request.

## References

- Altmann, J. (1974). Observational Study of Behavior: Sampling Methods. *Behaviour*, *49*(4), 227–267.
- Arcadi, A. C. (2000). Vocal responsiveness in male wild chimpanzees: Implications for the evolution of language. *Journal of Human Evolution*, *39*(2), 205–223. <https://doi.org/10.1006/jhev.2000.0415>
- Arlet, M., Jubin, R., Masataka, N., & Lemasson, A. (2015). Grooming-at-a-distance by exchanging calls in non-human primates. *Biology Letters*, *11*(10), 20150711–20150714. <https://doi.org/10.1098/rsbl.2015.0711>
- Bailey, W. J. (2003). Insect duets: Underlying mechanisms and their evolution. *Physiological Entomology*, *28*, 157–174. <https://doi.org/10.1046/j.1365-3032.2003.00337.x>
- Bascom, W. R. (1942). The principle of seniority in the social structure of the Yoruba. *American Anthropologist*, *44*, 37–46. <https://doi.org/10.1525/aa.1942.44.1.02a00050>
- Bouchet, H., Pellier, A. S., Blois-Heulin, C., & Lemasson, A. (2010). Sex differences in the vocal repertoire of adult red-capped mangabeys (*Cercocebus torquatus*): A multi-level acoustic analysis. *American Journal of Primatology*, *72*(4), 360–375. <https://doi.org/10.1002/ajp.20791>
- Cheney, D. L., Seyfarth, R. M., & Silk, J. B. (1995). The Responses of Female Baboons (*Papio cynocephalus ursinus*) to Anomalous Social Interactions: Evidence for Causal Reasoning? *Journal of Comparative Psychology*, *109*(2), 134–141. <https://doi.org/10.1037/0735-7036.109.2.134>
- Cutting, A. L., & Dunn, J. (2006). Conversations with siblings and with friends: Links between relationship quality and social understanding. *British Journal of Developmental Psychology*, *24*,

73–87. <https://doi.org/10.1348/026151005X70337>

- Da Cunha, R. G. ., & Byrne, R. W. (2006). Roars of Black Howler Monkeys (*Alouatta caraya*): Evidence for a Function in Inter-Group Spacing. *Behaviour*, *143*(10), 1169–1199.
- deWaal, F. B. M. (1982). *Chimpanzee Politics*. (Jonathan Cape, Ed.). London.
- Digweed, S. M., Fedigan, L. M., & Rendall, D. (2007). Who cares who calls? Selective responses to the lost calls of socially dominant group members in the white-faced capuchin (*Cebus capucinus*). *American Journal of Primatology*, *69*, 829–835. <https://doi.org/10.1002/ajp.20398>
- Dittus, W. P. J. (1984). Toque macaque food calls: Semantic communication concerning food distribution in the environment. *Animal Behaviour*, *32*, 470–477. [https://doi.org/10.1016/S0003-3472\(84\)80283-3](https://doi.org/10.1016/S0003-3472(84)80283-3)
- Duffy, K. G. (2006). *Social dynamics of male chimpanzees: adaptive significance of male bonds*. Los Angeles: University of California.
- Dunbar, R. I. M. (1996). *Grooming, gossip and the evolution of language*. London, United Kingdom: Faber & Faber.
- Dunbar, R. I. M. (2009). The social brain hypothesis and its implications for social evolution. *Annals of Human Biology*, *36*(5), 562–572. <https://doi.org/10.1080/03014460902960289>
- Fedurek, P., Machanda, Z. P., Schel, A. M., & Slocombe, K. E. (2013). Pant hoot chorusing and social bonds in male chimpanzees. *Animal Behaviour*, *86*(1), 189–196. <https://doi.org/10.1016/j.anbehav.2013.05.010>
- Fedurek, P., Slocombe, K. E., & Zuberbühler, K. (2015). Chimpanzees communicate to two different audiences during aggressive interactions. *Animal Behaviour*, *110*, 21–28. <https://doi.org/10.1016/j.anbehav.2015.09.010>
- Fischer, J., Kitchen, D. M., Seyfarth, R. M., & Cheney, D. L. (2004). Baboon loud calls advertise male quality: Acoustic features and their relation to rank, age, and exhaustion. *Behavioral Ecology and Sociobiology*, *56*, 140–148. <https://doi.org/10.1007/s00265-003-0739-4>
- Fröhlich, M., Kuchenbuch, P., Müller, G., Fruth, B., Furuichi, T., Wittig, R. M., & Pika, S. (2016). Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. *Scientific Reports*, *6*(25887), 1–14. <https://doi.org/10.1038/srep25887>
- Fuentes, A. (1998). Re-evaluating primate monogamy. *American Anthropologist*, *100*(4), 890–907.
- Funkhouser, J. A., Mayhew, J. A., & Mulcahy, J. B. (2018). Social network and dominance hierarchy analyses at Chimpanzee Sanctuary Northwest. *PlosOne*, *13*(2), e0191898. <https://doi.org/10.1371/journal.pone.0191898>
- Gautier, J.-P. (1988). Interspecific affinities among guenons as deduced from vocalizations. In *primate radiation: Evolutionary biology of the African guenons* (pp. 194–226). Cambridge: Cambridge University Press.
- Geissmann, T. (2002). Duet-splitting and the evolution of gibbon songs. *Biological Reviews of the Cambridge Philosophical Society*, *77*, 57–76. <https://doi.org/10.1017/S1464793101005826>
- Ghiglione, R. (1986). *L'homme communiquant*. (A. Colin, Ed.). Paris, France.
- Golinkoff, R. M. (1986). 'I beg your pardon?': The preverbal negotiation of failed messages. *Journal of Child Language*, *21*, 205–226. <https://doi.org/10.1017/S0305000900006826>
- Goodall, J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge: Belknap Press of Harvard University Press.
- Gouzoules, B. Y. S., Gouzoules, H., Marler, P., Rockefeller, T., & York, N. (1984a). Rhesus Monkey (

- Macaca Mulatta ) Screams : Agonistic Aid, 182–193.
- Gouzoules, S., Gouzoules, H., & Marler, P. (1984b). Rhesus Monkey (*Macaca mulata*) Screams: Representational Signaling in the Recruitment of Agonistic Aid. *Animal Behaviour*, *32*, 182–193.
- Graham, K. E., Hobaiter, C., Ounsley, J., Furuichi, T., & Byrne, R. W. (2018). Bonobo and chimpanzee gestures overlap extensively in meaning. *PLoS Biology*, *16*(2). <https://doi.org/10.1371/journal.pbio.2004825>
- Gros-Louis, J., & Mitani, J. (1998). Chorusing and call convergence in chimpanzees: Tests of three hypotheses. *Behaviour*, *135*(8), 1041–1064. <https://doi.org/10.1163/156853998792913483>
- Grubb, P., Butynski, T. M., Oates, J. F., Bearder, S. K., Disotell, T. R., Groves, C. P., & Struhsaker, T. T. (2003). Assessment of the Diversity of African Primates. *International Journal of Primatology*, *24*(6), 1301–1357.
- Gruber, T., & Clay, Z. (2016). A Comparison Between Bonobos and Chimpanzees: A Review and Update. *Evolutionary Anthropology*, *25*, 239–252. <https://doi.org/10.1002/evan.21501>
- Hammerschmidt, K., & Fischer, J. (2008). Constraints in primate vocal production. In O. K. Griebel U (Ed.), *The Evolution of Communicative Creativity: From Fixed Signals to Contextual Flexibility* (pp. 93–120). MIT Press.
- Harris, T. R., Fitch, W. T., Goldstein, L. M., & Fashing, P. J. (2006). Black and white colobus monkey (*Colobus guereza*) roars as a source of both honest and exaggerated information about body mass. *Ethology*, *112*(9), 911–920. <https://doi.org/10.1111/j.1439-0310.2006.01247.x>
- Hauser, M. D., & Marler, P. (1993). Food-associated calls in rhesus macaques (*Macaca mulatta*): I. Socioecological factors. *Behavioral Ecology*, *4*(3), 194–205. <https://doi.org/10.1093/beheco/4.3.194>
- Henry, L., Craig, A. J. F. K., Lemasson, A., & Hausberger, M. (2015). Social coordination in animal vocal interactions. Is there any evidence of turn-taking? The starling as an animal model. *Frontiers in Psychology*, *6*, 1416. <https://doi.org/10.3389/fpsyg.2015.01924>
- Kappeler, P. M., & van Schaik, C. P. (2002). Evolution of Primate Social Systems. *International Journal of Primatology*, *23*(4), 707–740.
- Koda, H. (2004). Flexibility and context-sensitivity during the vocal exchange of coo calls in wild Japanese macaques (*Macaca fuscata yakui*). *Behaviour*, *141*(10), 1279–1296. <https://doi.org/10.1163/1568539042729685>
- Langergraber, K., Mitani, J., & Vigilant, L. (2009). Kinship and social bonds in female chimpanzees (*Pan troglodytes*). *American Journal of Primatology*, *71*(10), 840–851. <https://doi.org/10.1002/ajp.20711>
- Langmore, N. E. (2002). Vocal duetting: definitions, discoveries and directions. *Trends in Ecology and Evolution*, *17*(10), 451–452.
- Laporte, M. N. C., & Zuberbühler, K. (2010). Vocal greeting behaviour in wild chimpanzee females. *Animal Behaviour*, *80*(3), 467–473. <https://doi.org/10.1016/j.anbehav.2010.06.005>
- Lehmann, J., & Boesch, C. (2009). Sociality of the dispersing sex: the nature of social bonds in West African female chimpanzees, *Pan troglodytes*. *Animal Behaviour*, *77*, 377–387. <https://doi.org/10.1016/j.anbehav.2008.09.038>
- Lemasson, A., Gandon, E., & Hausberger, M. (2010a). Attention to elders' voice in non-human primates. *Biology Letters*, *6*(3), 325–328. <https://doi.org/10.1098/rsbl.2009.0875>
- Lemasson, A., Ouattara, K., Bouchet, H., & Zuberbühler, K. (2010b). Speed of call delivery is related to context and caller identity in Campbell's monkey *mal*. *Naturwissenschaften*, *97*, 1–5.

- Lemasson, Alban, Guilloux, M., Rizaldi, Barbu, S., Lacroix, A., & Koda, H. (2013). Age- and sex-dependent contact call usage in Japanese macaques. *Primates*, *54*, 283–291. <https://doi.org/10.1007/s10329-013-0347-5>
- Lemasson, Alban, & Hausberger, M. (2004). Patterns of vocal sharing and social dynamics in a captive group of Campbell's monkeys (*Cercopithecus campbelli campbelli*). *Journal of Comparative Psychology*, *118*(3), 347–359. <https://doi.org/10.1037/0735-7036.118.3.347>
- Lemasson, Alban, Pereira, H., & Levréro, F. (2018). Social Basis of Vocal Interactions in Western Lowland Gorillas (*Gorilla g. gorilla*). *Journal of Comparative Psychology*, *132*(2), 141–151. <https://doi.org/10.1037/com0000105>
- Levinson, S. C. (2016). Turn-taking in Human Communication – Origins and Implications for Language Processing. *Trends in Cognitive Sciences*, *20*(1), 6–14. <https://doi.org/10.1016/j.tics.2015.10.010>
- Levréro, F., & Mathevon, N. (2013). Vocal Signature in Wild Infant Chimpanzees. *American Journal of Primatology*, *75*(4), 324–332. <https://doi.org/10.1002/ajp.22108>
- Levréro, F., Touitou, S., Frédet, J., Nairaud, B., Guéry, J.-P., & Lemasson, A. (2019). Social bonding drives vocal exchanges in Bonobos. *Scientific Reports*, *9*, 711. <https://doi.org/10.1038/s41598-018-36024-9>
- Manser, M. B. (2001). The acoustic structure of suricates' alarm calls varies with predator type and the level of response urgency. *Proceedings of the Royal Society B: Biological Sciences*, *268*(1483), 2315–2324. <https://doi.org/10.1098/rspb.2001.1773>
- Manser, M. B., Seyfarth, R. M., & Cheney, D. L. (2002). Suricate alarm calls signal predator class and urgency. *TRENDS in Cognitive Sciences*, *6*(2), 55–57.
- Masataka, N., & Biben, M. (1987). Temporal Rules Regulating Affiliative Vocal Exchanges of Squirrel Monkeys. *Behaviour*, *101*, 311–319.
- McComb, K., & Semple, S. (2005). Coevolution of vocal communication and sociality in primates. *Biology Letters*, *1*(4), 381–385. <https://doi.org/10.1098/rsbl.2005.0366>
- Mitani, J. C. (2009). Cooperation and competition in chimpanzees: Current understanding and future challenges. *Evolutionary Anthropology*, *18*, 215–227. <https://doi.org/10.1002/evan.20229>
- Mitani, J. C., & Gros-Louis, J. (1998). Chorusing and call convergence in chimpanzees: tests of three hypotheses. *Behaviour*, *135*(9–10), 1041–1064.
- Mitani, J. C., & Nishida, T. (1993). Contexts and social correlates of long-distance calling by male chimpanzees. *Animal Behaviour*, *45*(4), 735–746. <https://doi.org/10.1006/anbe.1993.1088>
- Muller, M. N., & Mitani, J. C. (2005). Conflict and Cooperation in Wild Chimpanzees. *Advances in the Study of Behavior*, *35*, 275–331. [https://doi.org/10.1016/S0065-3454\(05\)35007-8](https://doi.org/10.1016/S0065-3454(05)35007-8)
- Nishida, T. (1968). The social group of wild chimpanzees in the Mahali Mountains. *Primates*, *9*, 167–224. <https://doi.org/10.1007/BF01730971>
- Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009). Campbell's monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences*, *106*, 22026–22031. <https://doi.org/10.1073/pnas.0908118106>
- Owren, M. J., & Rendall, D. (1997). An Affect-Conditioning Model of Nonhuman Primate Vocal Signaling. In Owings et al. (Ed.), *Perspective in Ethology, Volume 12* (Plenum Press., pp. 299–346). New York.
- Papworth, S., Böse, A.-S., Barker, J., Schel, A.-M., & Zuberbühler, K. (2008). Male blue monkeys alarm call in response to danger experienced by others. *Biology Letters*, *4*, 472–475. <https://doi.org/10.1098/rsbl.2008.0299>



- Pika, S., Wilkinson, R., Kendrick, K. H., & Vernes, S. C. (2018). Taking turns: bridging the gap between human and animal communication. *Proceedings of the Royal Society B*, 285(1880), 1–9. <https://doi.org/10.1098/rspb.2018.0598>
- Pollick, A. S., & De Waal, F. B. M. (2007). Ape gestures and language evolution. *Proceedings of the National Academy of Sciences*, 104(19), 8184–8189. [www.pnas.org/cgi/content/full/](http://www.pnas.org/cgi/content/full/)
- Pougnault, L., Levréro, F., & Lemasson, A. (2020a). Conversation among primate species. In N. Masataka (Ed.), *The Origins of Language Revisited - Differentiation from Music and the Emergence of Neurodiversity and Autism* (p. in Press). Singapor: Springer.
- Pougnault, Loïc, Levréro, F., Mulot, B., & Lemasson, A. (2020b). Breaking conversational rules matters to captive gorillas: A playback experiment. *Scientific Reports*, 10(1), 6947. <https://doi.org/10.1038/s41598-020-63923-7>
- Prieur, J., Pika, S., Barbu, S., & Blois-Heulin, C. (2016). A multifactorial investigation of captive chimpanzees' intraspecific gestural laterality. *Animal Behaviour*, 116, 31–43. <https://doi.org/10.1016/j.anbehav.2016.03.024>
- Ravignani, A., Bowling, D., & Fitch, W. T. (2014). Chorusing, synchrony and the evolutionary functions of rhythm. *Frontiers in Psychology*, 5, 1118. <https://doi.org/10.3389/fpsyg.2014.01118>
- Reynolds, V. (2005). *The chimpanzees of the Budongo Forest*. Oxford: Oxford University Press.
- Roberts, S. G. B., & Roberts, A. I. (2019). Social and ecological complexity is associated with gestural repertoire size of wild chimpanzees. *Integrative Zoology*, n/a, n/a. <https://doi.org/10.1111/1749-4877.12423>
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, 50, 696–735. <https://doi.org/10.2307/412243>
- Schamberg, I., Cheney, D. L., Clay, Z., Hohmann, G., & Seyfarth, R. M. (2016). Call combinations, vocal exchanges and interparty movement in wild bonobos. *Animal Behaviour*, 122, 109–116. <https://doi.org/10.1016/j.anbehav.2016.10.003>
- Schel, A. M., & Zuberbühler, K. (2012). Dawn chorusing in guereza colobus monkeys. *Behavioral Ecology and Sociobiology*, 66(3), 361–373. <https://doi.org/10.1007/s00265-011-1282-3>
- Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, 40(1–2), 227–256. [www.elsevier.com/locate/specom](http://www.elsevier.com/locate/specom)
- Siegel, S., & Castellan, N. J. (1988). *Nonparametric statistics for the behavioral sciences (2nd ed.)*. New York, NY.
- Slocombe, K. E., & Zuberbühler, K. (2005). Agonistic Screams in Wild Chimpanzees (*Pan troglodytes schweinfurthii*) Vary as a Function of Social Role. *Journal of Comparative Psychology*, 119(1), 67–77.
- Slocombe, K., & Zuberbühler, K. (2010). Vocal Communication in Chimpanzees. In *The Mind of the Chimpanzee: Ecological and Experimental Perspectives*. Chicago & London: University of Chicago Press.
- Stern, D. N., Jaffe, J., Beebe, B., & Bennett, S. L. (1975). Vocalizing in unison and in alternation: Two modes of communication within the mother- infant dyad. *Annals of the New York Academy of Sciences*, 263(1), 89–100. <https://doi.org/10.1111/j.1749-6632.1975.tb41574.x>
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106(26), 10587–10592. <https://doi.org/10.1073/pnas.0903616106>
- Sugiura, H. (1993). Temporal and Acoustic Correlates in Vocal Exchange of Coo Calls in Japanese

- Macaques. *Behaviour*, 124, 3–4. <https://doi.org/10.1163/156853993X00588>
- Sugiura, H. (1998). Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Animal Behaviour*, 55, 673–687.
- Takahashi, D. Y., Narayanan, D. Z., & Ghazanfar, A. A. (2013). Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys. *Current Biology*, 23, 2162–2168. <https://doi.org/10.1016/j.cub.2013.09.005>
- van Lawick-Goodall, J. (1968). The Behaviour of Free-living Chimpanzees in the Gombe Stream Reserve. *Animal Behaviour Monographs*, 1(3), 165–311. [https://doi.org/10.1016/S0066-1856\(68\)80003-2](https://doi.org/10.1016/S0066-1856(68)80003-2)
- Wakefield, M. (2002). Female chimpanzee allogrooming behavior at Ngogo, Kibale National Park, Uganda. *American Journal of Physical Anthropology (Supplement)*, 34, 160–161.
- Wich, S. A., & De Vries, H. (2006). Male monkeys remember which group members have given alarm calls. *Proceedings of the Royal Society B: Biological Sciences*, 273(1587), 735–740. <https://doi.org/10.1098/rspb.2005.3320>
- Williams, J. M., Hsien-Yang, L., & Pusey, A. E. (2002). Costs and benefits of grouping for female chimpanzees at Gombe. In & L. F. M. C. Boesch, G. Hohmann (Ed.), *Behavioural diversity in chimpanzees and bonobos*. Cambridge: Cambridge University Press.
- Wilson, M. L., Hauser, M. D., & Wrangham, R. W. (2001). Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Animal Behaviour*, 61(6), 1203–1216. <https://doi.org/10.1006/anbe.2000.1706>
- Yoshida, S., & Okanoya, K. (2005). Evolution of turntaking: a bio-cognitive perspective. *Cognitive Studies*, 12(3), 153–165.
- Zumpe, D., & Michael, R. P. (1986). Dominance index: A simple measure of relative dominance status in primates. *American Journal of Primatology*, 10, 291–300. <https://doi.org/10.1002/ajp.1350100402>

## Electronic Supplementary Materials

**Table SI** | Call types used in this study based on the acoustic features given by Slocombe and Zuberbühler (2010).

Type of vocalization	Acoustic description	Percentage of call type in the dataset (n=516)
Pant-Hoot	Long-distance vocalization with four distinct phases: introduction phase (low frequency <i>hoo</i> calls), buildup phase (increasingly loud panted <i>hoo</i> calls with energy in both the inhalation and exhalation), climax phase (screams or roars) and letdown phase (resembles to buildup phase but with progressively decreasing energy). All four phases are not always including (introduction and letdown are commonly omitted).	11.6 %
Whimper	Soft low-frequency <i>hoo</i> calls. Whimpers are often produced with a pout expression.	26.6 %
Scream	Loud, high-pitched harmonic vocalizations with varying degrees of tonality. Usually, the acoustic energy is present only during exhalation.	25.3 %
Squeak	High-pitched, short calls. The calls are tonal signals often with clear harmonic structures.	4.1 %
Bark	Sharp, loud calls with abrupt onsets. They are often noisy and are generally low-pitched vocalizations.	1.4 %
†Waa-Bark	A distinct loud bark variant in which the sound <i>waa</i> is clear. Acoustically distinct from the shorter barks and screams. Waa-barks typically have a low frequency “w” introduction at call onset, and then a clear rise and short fall in pitch during the “aa” element of call.	0 %
Cough	Low-frequency vocalization, similar to grunt but rarely voiced. A short noisy signal with no harmonic structure.	8.3 %
Grunt	Short, soft low-frequency calls.	8.3 %
†Rough Grunt	Rough grunts grade from low-frequency, unvoiced, noisy grunts to high-frequency tonal squeaks.	0 %
Pant-Grunt	Noisy, low frequency grunts or barks, panted in a rapid rhythm with audible energy in both inhalation and exhalation.	1.4 %
†Pant	Unvoiced, soft, low-frequency sounds. Temporal patterning is regular and rapid.	0 %
†Huu	Tonal calls with most energy at onset and a rise and fall in frequency over the call. <i>Huu</i> calls can be loud.	0 %
Laugh	Noisy, low-frequency grunts and moans delivered in an irregular rhythm, reminiscent of hoarse wheezing human laughter. Acoustic energy is audibly present in both inhalation and exhalation, with most voicing occurring during inhalation.	7.0 %
‡Hoo	<i>Hoo</i> calls are tonal signals with a variable number of harmonics.	4.8 %

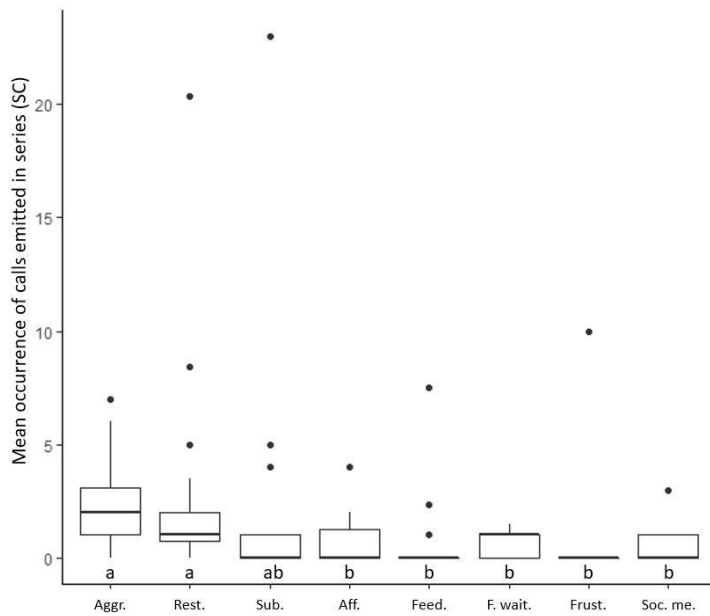
†Not recorded during our data collection period; ‡Recorded during data collection, described by Goodall (1986) but not by Slocombe and Zuberbühler (2010). The *mixed-call* represented 1.2 % of the total vocal production (Lemasson et al., 2018).

**Table SII** | Influence of social indices on the mean duration of call intervals and on the mean duration of calls in SC. For the dominance index, the social integration, the spatial proximity and the age of individuals, Spearman’s correlations were run, and the *rho* and *p-value* are given in the table ( $P > 0.05$ : NS;  $P < 0.05$ : \*;  $P < 0.01$ : \*\*;  $P < 0.001$ : \*\*\*). For the comparison regarding the sex of individuals, Wilcoxon rank sum tests were run and the table gives *W* and *p-value*.

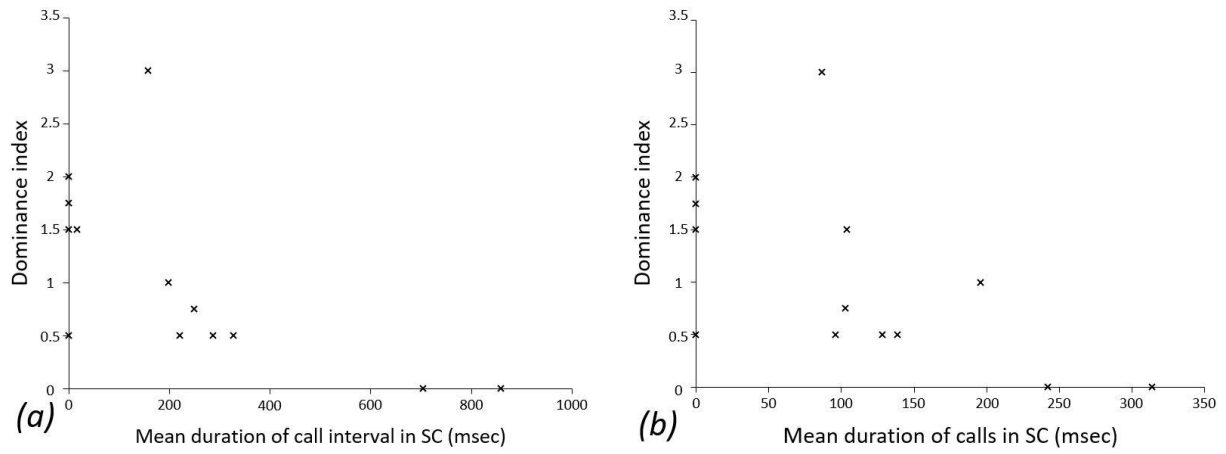
Mean duration of inter-call delay					
Contexts	Dominance Index	Social integration	Spatial proximity	Age	Sex
Affiliative	0.049 ; NS	-0.13 ; NS	-0.27 ; NS	-0.12 ; NS	25 ; NS
Aggressive	-0.75 ; **	-0.45 ; NS	-0.22 ; NS	0.089 ; NS	21 ; NS
Submissive	-0.29 ; NS	0.051 ; NS	0.27 ; NS	-0.26 ; NS	1 ; NS
Frustration	NA ; NA	0.37 ; NS	0.12 ; NS	-0.37 ; NS	22.5 ; NS
Social Meeting	-0.2 ; NS	0.43 ; NS	0.37 ; NS	-0.43 ; NS	30 ; NS
Feeding	-0.13 ; NS	0.19 ; NS	-0.024 ; NS	-0.37 ; NS	25 ; NS
Food Waiting	0.058 ; NS	0.36 ; NS	0.027 ; NS	-0.024 ; NS	25 ; NS
Resting	0.22 ; NS	-0.021 ; NS	-0.27 ; NS	0.064 ; NS	20.5 ; NS

Mean duration of consecutive calls					
Contexts	Dominance Index	Social integration	Spatial proximity	Age	Sex
Affiliative	-0.051 ; NS	-0.15 ; NS	-0.28 ; NS	-0.16 ; NS	26 ; NS
Aggressive	-0.66 ; *	-0.45 ; NS	-0.29 ; NS	0.17 ; NS	26 ; NS
Submissive	-0.29 ; NS	0.051 ; NS	0.27 ; NS	-0.26 ; NS	28.5 ; NS
Frustration	NA ; NA	0.37 ; NS	0.12 ; NS	-0.37 ; NS	22.5 ; NS
Social Meeting	-0.20 ; NS	0.43 ; NS	0.37 ; NS	-0.43 ; NS	30 ; NS
Feeding	-0.13 ; NS	0.19 ; NS	-0.024 ; NS	-0.37 ; NS	25 ; NS
Food Waiting	0.12 ; NS	0.37 ; NS	0.018 ; NS	0.024 ; NS	26 ; NS
Resting	0.32 ; NS	-0.085 ; NS	-0.27 ; NS	0.077 ; NS	19.5 ; NS



**Figure SI** | Mean occurrences of the calls produced in series (SC) in each context (Aggr.: Aggressive; Rest.: Resting; Sub.: Submissive; Aff.: Affiliative; Feed.: Feeding; F. wait.: Food waiting; Frust.: Frustration; Soc. me.: Social meeting). Box and whisker plots report the median, 25th and 75th percentiles, and the lowest and highest data, which were no more than 1.5 interquartile range from the box. When the letter indicated under the box plots is the same there was no significant difference between them (Generalized Linear Mixed Model with a negative binomial distribution).



**Figure SII** | (a) *Dominance index in relation to the mean duration of call interval in SC (Spearman correlation,  $P < 0.01$ ), and (b) to the mean duration of the calls in SC in aggressive context (Spearman correlation, 0.05).*



# Chapitre V

## **Influence du système social sur l'émergence des échanges vocaux organisés**







# Article 4

## **“Conversational rules” are driven by social pressures, say studies on great apes**

Loïc Pournault<sup>1,2,7</sup>, Florence Levréro<sup>2</sup>, Maël Leroux<sup>3,4,5</sup>, Julien Paulet<sup>1</sup>, Pablo Bombani<sup>6</sup>, Fabrice Dentressangle<sup>2</sup>, Laure Deruti<sup>2</sup>, Baptiste Mulot<sup>7</sup>, Alban Lemasson<sup>1,8</sup>

<sup>1</sup> Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000 Rennes, France.

<sup>2</sup> Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES / CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France.

<sup>3</sup> Department of Comparative Linguistics, University of Zürich, Zürich, Switzerland.

<sup>4</sup> Budongo Conservation Field Station, Masindi, Uganda.

<sup>5</sup> Center for the Interdisciplinary Study of Language Evolution (ISLE), University of Zürich, Zürich, Switzerland.

<sup>6</sup> NGO Mbou-Mon-Tour (Republic Democratic of Congo).

<sup>7</sup> ZooParc de Beauval & Beauval Nature, 41110 Saint Aignan, France.

<sup>8</sup> Institut Universitaire de France, Paris, France.

*Article en préparation*



## Synthèse : article 4



Que nous apprennent les études sur les grands singes quant aux pressions sociales sous-jacentes à l'émergence d'échanges vocaux « conversationnels » ?

### Contexte

L'hypothèse selon laquelle les espèces ayant un système social complexe exprimeraient une communication complexe prend de l'ampleur depuis plusieurs dizaines d'années. Néanmoins, bien que les relations inter-individuelles sont au cœur de la définition de la socialité, les critères traditionnellement utilisés pour décrire la complexité de la communication font souvent fit de sa dimension interactionnelle. Pour certains auteurs néanmoins, la communication ne se réduit pas à ce qui peut être exprimé par un individu, mais aussi à la manière dont les interactions vocales se structurent. Ainsi, de nombreuses études ont révélé l'existence d'interactions vocales régies par des règles aussi bien temporelles que sociales chez les primates non-humains, soulignant un possible parallèle avec les conversations humaines. Cependant, bien que seules des études comparatives permettraient d'aider à démêler le scénario évolutif de cette capacité langagière, les études empiriques se focalisent sur les primates non-hominoïdes, délaissant les grands singes. Ces derniers représentent pourtant un taxon particulièrement intéressant pour de telles études comparatives en raison de leur proximité phylogénétique avec l'humain et de la grande diversité de leurs systèmes sociaux. Au cours de cette étude, nous avons cherché à comprendre comment les besoins de la vie sociale peuvent avoir influencés l'émergence de certaines formes d'interactions vocales entre membres d'un même groupe chez quatre espèces de grands singes que sont les chimpanzés, les bonobos, les gorilles des plaines de l'Ouest et les orangs-outans de Bornéo.

### Méthode

#### ① Synthèse de la littérature

Résultats provenant de la littérature contemporaine ayant attiré à la socio-écologie et la communication chez les grands singes (*i.e.* bonobos, chimpanzés, gorilles des plaines de l'Ouest et orangs-outans de Bornéo)

#### ② Collectes des données vocales (si absente de la littérature)

Captivité		Milieu sauvage	
Orangs-outans		Chimpanzés (N= 35)	Bonobos (N=16)
<i>Toutes occurrences</i> (Sessions de 1hr)		<i>Toutes occurrences &amp; Ad libitum</i> (Sessions de 2hr)	<i>Ad libitum</i> Total : 91hr04
ZooParc de Beauval (N=8) Total : 17hr	Ouwehand DierenPark (N=8) Total : 25hr	Total : 44hr34	

#### ③ Descriptions temporelles des productions vocales

- A l'échelle du groupe :  
Emises en interaction VS hors interaction
- A l'échelle de l'individu :  
Emises : de façon unique VS répétée

### Principaux résultats

- Seule l'espèce la moins socialement complexe, les orangs-outans, ne produit pas d'interactions vocales.
- La proportion de vocalisations superposées est plus élevée chez l'espèce la plus compétitrice, les chimpanzés.
- Les échanges vocaux organisés, qui respectent des tours de parole, sont majoritaires chez les espèces les plus tolérantes, les bonobos et les gorilles.
- Les types de cris utilisés et les principaux interlocuteurs ne sont pas choisis aléatoirement.

### Conclusion

Notre étude met en lumière l'existence d'échanges vocaux « conversationnels » chez les grands singes, rendant plausible l'idée d'une compétence ancestrale ancrée profondément dans la lignée des primates. Par ailleurs, la forme prise par les expressions vocales semble correspondre à la qualité des liens sociaux existant entre les membres du groupe, supportant les théories d'une évolution guidée socialement. Enfin, l'approche comparative peut permettre d'identifier les multiples facteurs sociaux qui sous-tendent la complexité sociale et qui peuvent être à l'origine de l'évolution de la communication vocale.

## “Conversational rules” are driven by social pressures, say studies on great apes.

Loïc Pougault<sup>1,2,7</sup>, Florence Levréro<sup>2\*</sup>, Maël Leroux<sup>3,4,5</sup>, Julien Paulet<sup>1</sup>, Pablo Bombani<sup>6</sup>, Fabrice Dentressangle<sup>2</sup>, Laure Deruti<sup>2</sup>, Baptiste Mulot<sup>7</sup>, Alban Lemasson<sup>1,8\*</sup>

<sup>1</sup> Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000 Rennes, France.

<sup>2</sup> Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES / CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France.

<sup>3</sup> Department of Comparative Linguistics, University of Zürich, Zürich, Switzerland.

<sup>4</sup> Budongo Conservation Field Station, Masindi, Uganda.

<sup>5</sup> Center for the Interdisciplinary Study of Language Evolution (ISLE), University of Zürich, Zürich, Switzerland.

<sup>6</sup> NGO Mbou-Mon-Tour (Republic Democratic of Congo)

<sup>7</sup> ZooParc de Beauval & Beauval Nature, 41110 Saint Aignan, France.

<sup>8</sup> Institut Universitaire de France, Paris, France.

\* equal contribution

### Abstract

In the last decade, two theories, one on the evolution of animal vocal communication in general and the other on the origins of human language, have gained ground. The first theory argues that the complexity of communication co-evolves with the complexity of sociality. Species forming larger groups with complex social networks would have more elaborate vocal repertoires. The second theory posits that the core of communication is represented not only by what can be expressed by an isolated caller, but also, by the way that vocal interactions are structured, language being above all a social act. Primitive forms of conversational rules based on a vocal turn-taking principle are thought to exist in our primate cousins. Nonetheless, empirical data are lacking and comparative studies can help to untangle the evolutionary scenario. However, the majority of available studies on the topic focus on monkeys, primates that are relatively distant from the human genera. Great apes represent excellent candidates to run such comparative studies because of their phylogenetic proximity with humans and their varied lives. We propose that studying vocal turn-taking in apes can address several major gaps regarding the social relevance of vocal turn-taking and the evolutionary trajectory of this vocal behaviour among anthropoids.

Here, we investigated how the social needs of a species may influence the vocal interaction patterns observed among group members. We gathered data from the literature as well as original unpublished data (when lacking in the literature) on four great ape species: chimpanzees, bonobos, Western lowland gorillas and Bornean orang-utans.

We found that no clear-cut relationship between classical social complexity (*e.g.* number of group members, interaction rates) and vocal complexity parameters (*e.g.* vocal repertoire size, call rates), the nature of the primate society (*i.e.* group composition, diversity and valence of social bonds) and the type of vocal interaction patterns (isolated calling, call overlap, turn-taking based vocal exchanges) do appear somehow related. Namely, isolated calling is the main vocal pattern found in the less socially complex species (orang-utan), while the three other species show vocal interactions that are structured in a temporally-ruled way. A high proportion of overlapping vocalisations is found in the most competitive species (chimpanzee), while vocal turn-taking predominate in more tolerant bonobos and gorillas. Also, preferentially interacting individuals and call types used to interact are not randomly distributed. Indeed, vocal overlap

(“chorusing”) and vocal exchange (“conversing”) appear as possible social strategies used to advertise/strengthen social bonds.

In summary, our analyses highlight that: (1) vocal turn-taking is also observed in great apes, thus revealing universal rules for conversing that may be deeply rooted in the primate lineage, (2) vocal interaction patterns match the quality of social bonds supporting theories of a socially-driven evolution, (3) adopting a targeted comparative approach can help to identify the multiple and subtle social factors underlying social complexity and that may drive the evolution of vocal communication. We believe that focusing more on the coevolution between “social life style” and “vocal interaction patterns” can form the bases of a promising field of investigation.

**Keywords (5-10):** Communication evolution, Vocal communication, Social communication, Social structure, Social organization, Vocal exchange, Conversation, Turn-taking, Primates, Hominidae.

## CONTENTS

1. Introduction.....	166
1.1. Vocal communication abilities as a mirror of social complexity .....	166
1.2. Temporal calling patterns: the abundance and diversity of vocal interactions in primates .....	168
1.3. Vocal exchange in primates: a social function .....	171
1.4. The evolution of vocal turn-taking systems: an open question .....	173
2. Great apes’ socio ecology: similarities and differences.....	175
2.1. Origins and geographical repartition .....	181
2.2. Social systems.....	181
3. Vocal interactions in great apes .....	183
3.1. Repertoire size and hourly rate of vocal production.....	183
3.2. Occurrences and features of calling patterns across great ape species .....	184
3.3. Social features of vocal interactions .....	187
4. Conclusion .....	190

## 1. Introduction

### 1.1. Vocal communication abilities as a mirror of social complexity

Possibly because communication is widespread across the animal kingdom, and occurs in a high diversity of social contexts such as territorial defence, mate attraction, intergroup competition, intragroup socio-spatial coordination and social bonding (Waser, 1977; McComb, 1991; Geissmann, 1999; Arlet *et al.*, 2015), the evolutionary forces driving communication complexity are still intensively debated, arguably including sound propagation constraints, degree of visibility in the habitat, predation nature and rate, food chain, and life-history traits (*e.g.* Lemasson, 2011; Freeberg, Dunbar, & Ord, 2012; Fischer *et al.*, 2017a; Hammerschmidt & Fischer, 2019; Pisanski *et al.*, 2016).

The hypothesis that social complexity has been one of the driving forces shaping vocal complexity (if not *the* main driving force when it comes to primates) has persisted from Darwin's time (1877) to the present day (Freeberg *et al.*, 2012; Bouchet, Blois-Heulin, & Lemasson, 2013; Sewall, 2015). Nonetheless, the methodology typically used to investigate this hypothesis has been criticized notably for its lack of precision (Freeberg *et al.*, 2012; Fischer, Wadewitz, & Hammerschmidt, 2017b). One of the biggest challenges of this “social complexity hypothesis for communication” is the lack of clarity on what “complexity” means in terms of both social and communicative aspects (Manser *et al.*, 2014).

Group size was used as a proxy for social complexity in some of the earlier studies (Dunbar, 1998; McComb & Semple, 2005). For example, in the non-human primate phylogeny (NHP), group size appears to be correlated with the number of call types in the species' vocal repertoire (McComb & Semple, 2005). Group-living animals do indeed need to deal with a wide range of interactions with diverse social partners, and to transmit without ambiguity various emotional and motivational states (Morton, 1977; Freeberg *et al.*, 2012 but see Keenan *et al.*, 2020). More recently, authors pointed out the importance of also taking into account the

group's social structure (*e.g.* unit size, unit density and age-sex composition) and social organization (*e.g.* member roles, social bonds strength) to better quantify the degree of complexity of a social system (Lemasson, 2011; Freeberg *et al.*, 2012; Pollard & Blumstein, 2012; Freeberg & Krams, 2015; Fischer *et al.*, 2017a). In primates, for instance, it was found that the stability of male-female social bonds may be responsible for the emergence of derived call types during evolution (see an example comparing baboons' vocal repertoires: Gustison, le Roux, & Bergman, 2012).

The most common way to measure communicative complexity within a signalling system of a particular species is the number of structurally and functionally distinct elements and/or the amount of bits of information in a call (Freeberg *et al.*, 2012). Interestingly, the different call types of a given vocal repertoire are not equally complex. "Social" calls (such as contact calls) are more individually variable than alarm calls, for instance (Lemasson & Hausberger, 2011; Bouchet *et al.*, 2013). Nonetheless, it is not always easy to quantify communicative complexity using these features. Very often, signals exist in graded form, and without clearly identifying the acoustic boundaries between signal types, we may have difficulties to interpret their true functions (Keenan, Lemasson, & Zuberbühler, 2013; Fischer, Wadewitz, & Hammerschmidt, 2017b, Keenan *et al.*, 2020). Also, the perception/recognition of the amount of bits of information is based on the assumption that the higher the variability of elements in a signalling system, the greater the potential for information or complexity within the system (Freeberg *et al.*, 2012). A higher rate of information would thus provide more meaning, that implicitly suggests that the species have the cognitive abilities to process this information (Manser *et al.*, 2014). Yet, to make a conclusion on the complexity of a communicative system, it is therefore necessary to experimentally test how meaningful the observed variation in a signal is to receivers (Manser *et al.*, 2014; Garcia *et al.*, 2020).



Thus, the current methodology used to argue in favour of the “social complexity hypothesis for communication” remains unsatisfactory and comparative approaches with even more multi-factorial analyses are urgently needed. Recently, Rebout and collaborators (2020) studied two components of vocal complexity (diversity - number of categories of calls, and flexibility - degree of gradation between categories of calls) and compared the acoustic structure of vocal signals in groups of macaques belonging to four species with varying levels of uncertainty (i.e. social complexity) in social tolerance (i.e. the higher the degree of tolerance, the higher the degree of uncertainty). The results showed that tolerant macaques displayed higher levels of vocal diversity and flexibility than did intolerant macaques in situations with a greater number of options and consequences, i.e. in agonistic and affiliative contexts.

Communicating and socializing are above all interacting acts. To explore the link between the requirements/characteristics of social life and the expression/structure of vocal exchanges is another way to apprehend the “social complexity hypothesis for communication”. All the aforementioned studies and theories focused on vocal repertoire sizes and acoustic structure variations, with thus a focus on individuals as if they were isolated callers. However, the social core of communication may also be found in the patterns and rules of vocal interactions among two or more individuals (Lemasson, Hausberger, & Zuberbühler, 2005; Lemasson, 2011; Fedurek *et al.*, 2013a; Arlet *et al.*, 2015; Henry *et al.*, 2015; Pika, Wilkinson, Kendrick & Vernes, 2018). Focusing on the “social life style – vocal interaction pattern” coevolution offers a promising new field of investigation.

## 1.2. Temporal calling patterns: the abundance and diversity of vocal interactions in primates

When focusing on the temporal rules guiding call production in animals, three main types of temporal calling patterns are typically found (Pougnault, Levréro, & Lemasson,

2020a): (i) isolated call (when only one individual produces a single call), (ii) series of consecutive calls (a given call type is repeated several times in a row by a given individual), and (iii) vocal interaction (when at least two individuals temporally coordinate their vocal productions). The usage rate of these different calling patterns may depend on the context and call type, but also on the identity (*i.e.* age, sex) and social characteristics of the caller.

Vocal interactions are widespread and found across various clades, from insects (e.g. Greenfield, 1994; Drosopoulos & Claridge, 2005) to vertebrates such as anurans (e.g. Grafe, 1996, 1999), birds (e.g. Naguib & Kipper, 2006; Naguib & Mennill, 2010), and mammals (e.g. non-primate species: elephants *Loxodonta africana* Leong *et al.*, 2003; sperm-whales *Physeter macrocephalus* Schulz *et al.*, 2008; common bottlenose dolphins *Tursiops truncatus* Janik, 2000; naked mole-rats *Heterocephalus glaber* Yoshida & Okanoya, 2005; bats *Diaemus youngi* Carter *et al.*, 2008), including a broad range of non-human primate species (e.g. Lemasson, Gandon, & Hausberger, 2010; Takahashi, Narayanan, & Ghazanfar, 2013; Pika *et al.*, 2018; Pougnauld *et al.*, 2020a). But what motivates and structures the organisation of these vocal interactions still remains an open question.

Because non-human primates show a large diversity of social systems with frequent vocal interactions regardless of their habitats (e.g. Kappeler & Van Schaik, 2002; Shultz & Dunbar, 2007), primates are ideal models for comparative studies. Primatologists typically distinguish several vocal interaction patterns. The disorganized phonoresponses occur when one individual produces a call, typically an alarm call that triggers calls in an apparent chaotic way from the other group members (e.g. in blue monkeys *Cercopithecus mitis stuhlmanni* Papworth *et al.*, 2008). In contrast, three main types of temporally organized vocal interactions are usually described: (i) choruses, which result from a simultaneous production of the same call type implying causal interactions between individuals joining the synchronous production (Yoshida & Okanoya, 2005; Ravignani, Bowling, & Fitch, 2014); (ii) duets, when two

individuals synchronize or overlap long series of calls or songs with a stereotyped temporal association (e.g. gibbons *Hylobates syndactylus* Geissmann & Orgeldinger, 2000), and (iii) vocal exchanges involving individuals who alternate their calling and vocally respond to one another.

Vocal exchanges are based on a turn-taking principle, and have been documented in all main primate branches (e.g. *Catarrhini*: Campbell's monkeys *Cercopithecus campbelli* Lemasson *et al.*, 2011a, Japanese macaques *Macaca fuscata* Sugiura, 1993; *Platyrrhini*: squirrel monkeys *Saimiri sciureus* Masataka & Biben, 1987, pygmy marmoset *Cebuella pygmaea* Snowdon & Elowson, 1999, spider monkeys *Ateles geoffroyi* Briseño-Jaramillo *et al.*, 2018; for a review, see Pougnault *et al.*, 2020a). Turn-taking is defined as the orderly exchange of communicative signals between individuals adjusting the timing of their vocal responses by respecting a minimum silence gap (preventing call overlap), and a maximum silence gap (allowing call coordination) (for a review see Pika *et al.*, 2018). Vocal exchanges in non-human primates are typically composed of one to three turns with individuals alternating their calls (e.g. Lemasson *et al.*, 2010; Levréro *et al.*, 2019), and thus are typically characterised by short vocal sequences. The common response delay is rarely shorter than the average call duration, thus preventing call overlap (e.g. Diana's monkeys *Cercopithecus diana*: Candiotti, Zuberbuhler, & Lemasson, 2012; squirrel monkeys: Masataka & Biben, 1987; Japanese macaques: Sugiura, 1993; spider monkeys: Briseño-Jaramillo *et al.*, 2018; gorillas *Gorilla gorilla gorilla*: Lemasson, Pereira & Levréro, 2018; bonobos *Pan paniscus*: Levréro *et al.*, 2019). Interestingly, speech overlap avoidance and turn-taking also represent the social core of human vocal interactions. It is found in all human cultures (Stivers *et al.*, 2009; Levinson, 2016), while choruses and duets are absent except when humans sing (Welch, 2005).

Beyond these temporal aspects, the turn-taking based vocal exchanges of non-human primates differ from the other types of vocal interactions regarding several features. First, they

involve a diversity of potential recurrent vocal partners, which differs from duets, where the possible partners are generally limited (MacKinnon & MacKinnon, 1980; Geissmann & Orgeldinger, 2000). Second, vocal exchange interlocutors can be of any age or either sex (which differs from most choruses: Haimoff, 1986). Third, vocal exchanges are not restricted to a specific context, which differs from – usually long-distance – collective communication associated with mate attraction, territory protection and environmental disturbance (Briseño-Jaramillo *et al.*, 2017), and are also not restricted to a specific time of the day or year (as opposed to morning choruses and reproductive signals Gros-Louis & Mitani, 1998; Schel & Zuberbühler, 2012). Most commonly, these vocal exchanges are observed daily in short-distance communication, and typically occur throughout the entire day. In most cases, vocal exchanges are soft calls exchanged in peaceful and relaxed contexts (Pougnault *et al.*, 2020a).

### 1.3. Vocal exchange in primates: a social function

The “grooming-at-distance” hypothesis proposed by Dunbar (1996) highlighted the key role played by vocal exchanges during the evolution of primates. This theory suggested that, in large social groups and in groups with solid and long-lasting bonds, the function of gestural grooming was partially transferred to vocal exchanges during evolution. Gestural grooming plays an important role in forming new bonds and strengthening existing ones in all primate species (Dunbar, 1991). As group size and bonding network increased during our ancestral history, vocal exchanges may have developed, as it ultimately would have become impossible to allocate sufficient time to groom all partners physically.

Vocal exchanges do indeed occur mostly between interlocutors that are non-randomly chosen, depending on, according to species, their affiliative bonds (e.g. bonobos (Levréro *et al.*, 2019), spider monkeys *Ateles geoffroyi* Briseño-Jaramillo *et al.*, 2018, Japanese macaques Arlet *et al.*, 2015, pygmy marmosets Snowdon & Cleveland, 1984,) and/or respective ages (common

marmosets *Callithrix jacchus* Chen, Kaplan, & Rogers, 2009, Campbell's monkeys Lemasson *et al.*, 2010, for a review Pougnault *et al.*, 2020a). The temporal organization of vocal exchanges in monkeys is also flexible and often socially-guided. The inter-call delay during vocal turn-taking can be shorter when interlocutors have strong affiliative bonds (squirrel monkeys Biben Symmes & Masataka, 1986; Masataka & Biben, 1987). The phenomenon of vocal convergence is also reported in monkeys: over more or less long periods of time, individuals can modify the frequency modulation pattern of their contact call to copy their preferred partners (*e.g.* Campbell monkeys: Lemasson & Hausberger, 2004; pygmy marmosets: Snowdon & Elowson, 1999; Lemasson *et al.*, 2011b for a review) or even match the acoustic structure of their vocal response with the acoustic structure of the call they are responding too (Japanese macaques: Sugiura & Masataka, 1995; Diana monkeys: Candiotti, Zuberbühler, & Lemasson, 2012; spider monkey: Briseño-Jaramillo *et al.*, 2018). These vocal exchanges are definitely more than just emotional discharges because they require a significant cognitive demand implying a certain degree of cooperation, which is reflected in the social synchronisation and in temporal planning to avoid vocal overlap while insuring the fluidity of the interaction (Takahashi *et al.*, 2013; Levinson & Holler, 2014; Levinson, 2016).

However, despite the growing interest in the matter, many gaps remain to be filled to fully apprehend the social relevance of vocal exchanges in non-human primates and the evolutionary trajectory of these types of vocal interactions. One of the major lacking elements is the weak representation of comparative studies that are necessary to characterise the vocal exchange phenotype of a wide diversity of primate species. This could be achieved through systematic testing, which should then enable researchers to map out phylogenetically the evolution of primates' vocal exchanges (Pika *et al.*, 2018). Moreover, the use of experimental procedures, such as playback experiments that allow the assessment of the importance of the

temporal and social organisation of vocal exchanges for non-human primates, is also sorely needed.

#### 1.4. The evolution of vocal turn-taking systems: an open question

Several parallels between non-human animal vocal exchanges and human conversations have been observed (Snowdon & Cleveland, 1984; Sugiura & Masataka, 1995; Lemasson *et al.*, 2011b; Henry *et al.*, 2015). Overlap avoidance, turn-taking, social selection of interlocutors, varied contexts of occurrences going beyond a strict message transmission function and acoustic convergence are indeed found in both animal vocal exchanges and human conversations. Researchers have thus questioned the possible biological bases of human conversations and the role played by social factors during the evolution of vocal exchanges, notably by conducting cross-species comparisons.

It is worth noting that some conversational characteristics have been found in highly social non-primate species (*e.g.* elephants Leong *et al.*, 2003; sperm-whales Schulz *et al.*, 2008; common bottlenose dolphins Janik, 2000; starlings Henry *et al.*, 2015; naked mole-rats Yoshida & Okanoya, 2005; bats Carter *et al.*, 2008). Nonetheless, the strong interest allotted to non-human primates communication has made it possible to highlight numerous criteria allowing to consider their vocal exchanges as “conversation-like” (Pougnault *et al.*, 2020a).

However, most published research on vocal interactions has focused on *Platyrrhini* and *Catarrhini* monkeys or lesser apes (*e.g.* agile gibbons *Hylobates agilis* Koda *et al.*, 2013; siamangs *Hylobates syndactylus* Geissmann & Orgeldinger, 2000). Remarkably, researchers have largely neglected great apes for decades, even if they are also excellent candidates to display such vocal exchanges given their well described complex social life with long-term bonds and intricated social interactions (*e.g.* orang-utans *Pongo pygmaeus* Delgado & Van Schaik, 2000; gorillas Stokes, 2004; chimpanzees van Lawick-Goodall, 1968, Goodall, 1986;

bonobos Kano, 1982, 1992), and despite the central role of vocal communication in their daily lives (e.g. orang-utans Hardus *et al.*, 2009; gorillas Salmi, Hammerschmidt, & Doran-Sheehy, 2013; chimpanzees Slocombe & Zuberbühler, 2010; bonobos De Waal, 1988, Bermejo & Omedes, 1999), and their close phylogenetic proximity to the human species. As a result, the selective pressures that maintain such vocal behaviours in non-human primates remain debated, notably because the current state of knowledge still prevents to conclude in favour of a convergent evolution scenario (guided in some species by the requirements of social life) or a shared inheritance (suggesting an ancient mechanism present deep inside the primate lineage).

Another ongoing debate is whether these abilities to converse in animals are genetically programmed or, as in humans (Stern *et al.*, 1975; Ginsburg & Kilbourne, 1988), socially acquired. The few available studies on monkeys showed that juveniles break conversational rules more often than do adults (howler monkeys *Alouatta pigra* Briseño Jaramillo *et al.*, 2017; Campbell's monkeys Lemasson *et al.*, 2011a). A recent study conducted on marmosets suggested that, rather than parental influence (Chow, Mitchell, & Miller, 2015), the ability to respect vocal turn-taking might be due to self-monitoring which increases with age (Takahashi, Fenley, & Ghazanfar, 2016). Despite the lack of empirical evidence to answer this question, results from playback experiments using a violation-of-expectation paradigm showed that adult monkeys clearly discriminated between conversationally appropriate and inappropriate vocal exchange patterns, whereas socially inexperienced juveniles did not (Japanese macaques Bouchet, Koda, & Lemasson, 2017; Campbell's monkeys Lemasson *et al.*, 2011a). These findings showed that the audience detects the violation of expected social rules in vocal exchanges between third parties (Cheney, Seyfarth, & Silk, 1995; Wittig *et al.*, 2007; Lemasson *et al.*, 2011a; Bouchet *et al.*, 2017; Pournault *et al.*, 2020b), suggesting that the social rules in vocal interactions are not a simple neurobiologically guided behaviour but rather a social awareness of the respect of social rules.

Throughout this paper, we aim at questioning the possible *scenario* of social-vocal co-evolution, namely defining how the needs of social life may influence vocal interaction patterns among group members. The four species of great apes are valuable candidates to run a comparative analysis of social and vocal features because they present four very different and well-known social systems. We gathered data from the literature as well as original unpublished data when the vocal feature of interest was lacking in the literature for a given species (details of methods and results for the new data are available as *Supporting information 1, 2 and 3*).

## 2. Great apes' socio ecology: similarities and differences

Current knowledge on the socio-ecology of great ape species, particularly in terms of vocal communication, varies by species. We thus selected the following species/sub-species for which the literature is rich enough to run a proper comparative analysis: (i) bonobos (*Pan paniscus*), (ii) chimpanzees (*Pan troglodytes*), (iii) Western lowland gorilla (*Gorilla gorilla gorilla*), and (iv) Bornean orang-utans (*Pongo pygmaeus*). All key phylogenetic and socio-ecological references used to prepare the paragraphs below can be found in Table 1.



**Table 1|** Socio-ecology and vocal behaviour of four great ape species

		<b>Bornean orangutans</b> ( <i>Pongo pygmaeus</i> )	<b>Western lowland gorillas</b> ( <i>Gorilla gorilla gorilla</i> )	<b>Chimpanzees</b> ( <i>Pan troglodytes</i> )	<b>Bonobos</b> ( <i>Pan paniscus</i> )
<b>Last common ancestor with human species</b>		~9 Myr to ~18 Myr  (Hobolth <i>et al.</i> , 2011; Castellano & Munch, 2020)	~8.6 to ~11.5 Myr  (Upham, Esselstyn, & Jetz, 2019; Castellano & Munch, 2020)	~4.2 to ~9 Myr  (Patterson <i>et al.</i> , 2006; Hobolth <i>et al.</i> , 2011; Prüfer <i>et al.</i> , 2012; Rogers & Gibbs, 2014; Upham <i>et al.</i> , 2019)	~4.2 to ~9 Myr  (Patterson <i>et al.</i> , 2006; Hobolth <i>et al.</i> , 2011; Prüfer <i>et al.</i> , 2012; Rogers & Gibbs, 2014; Upham <i>et al.</i> , 2019)
<b>Socio-ecological characteristics</b>	<b>Geographical distribution</b>	South-East Equatorial Asia  (Ancrenaz <i>et al.</i> , 2018)	Western Equatorial Africa  (Maisels <i>et al.</i> , 2018)	Western, central and eastern Africa  (Oates <i>et al.</i> , 2007)	Central Africa  (Kano, 1992; Idani <i>et al.</i> , 2008; Fruth <i>et al.</i> , 2016)
	<b>Main habitat type</b>	Humid forest  (Delgado & van Schaik, 2000; Ancrenaz <i>et al.</i> , 2018)	Humid and <i>terra firma</i> lowland forests  (Gatti <i>et al.</i> , 2004; Levréro <i>et al.</i> , 2006; Maisels <i>et al.</i> , 2018; Forcina <i>et al.</i> , 2019)	Humid forest (sometimes savannah woodland)  (Nishida, 1968; van Lawick-Goodall, 1968; Goodall, 1986; Oates <i>et al.</i> , 2007)	Humid forest (sometimes forest-savannah mosaics)  (Kano, 1982; Narat <i>et al.</i> , 2015; Fruth <i>et al.</i> , 2016)
	<b>Group size</b>	Small (1, rarely 2 adults)  (Delgado & van Schaik, 2000; van Noordwijk <i>et al.</i> , 2012)	Medium (10, occasionally up to 20)  (Williamson & Butynski, 2013)	Large (20 to 120)  (Muller & Mitani, 2005)	Large (30 to 80)  (Fruth, Williamson, & Richardson, 2013; Fruth <i>et al.</i> , 2016)
	<b>Social group pattern</b>	Single ♂ / single ♀ with offspring (both sexes sometimes gregarious according to food availability)	Uni ♂ / Multi ♀ (temporary grouping composed mainly of immature ♂)  (Gatti <i>et al.</i> , 2004; Robbins <i>et al.</i> , 2004; Levréro <i>et al.</i> , 2006; Harcourt & Stewart, 2007)	Multi ♂ / Multi ♀ (Fission/Fusion society)  (Wrangham & Smuts, 1980; Muller & Mitani, 2005)	Multi ♂ / Multi ♀ (Fission/Fusion society)  (Kano, 1982; van Elsacker, Vervaecke, & Verheyen, 1995; Furuichi <i>et al.</i> , 2012)

		(Delgado & van Schaik, 2000; van Noordwijk <i>et al.</i> , 2012)			
<b>Sex dispersal</b>	♂	Disperse (Morrogh-Bernard <i>et al.</i> , 2011; van Noordwijk <i>et al.</i> , 2012)	Disperse (Stokes, Parnell, & Olejniczak, 2003; Forcina <i>et al.</i> , 2019)	Philopatric (Nishida, 1968; Wrangham & Smuts, 1980)	Philopatric (Kano, 1982; van Elsacker <i>et al.</i> , 1995; Furuichi <i>et al.</i> , 2012)
	♀	Philopatric (Arora <i>et al.</i> , 2010; Morrogh-Bernard <i>et al.</i> , 2011; Nater <i>et al.</i> , 2011)	Disperse (Stokes <i>et al.</i> , 2003; Forcina <i>et al.</i> , 2019)	Disperse (Nishida, 1968; Wrangham & Smuts, 1980)	Disperse (Kano, 1982; van Elsacker <i>et al.</i> , 1995; Furuichi <i>et al.</i> , 2012)
<b>Social bonds</b>	♂-♂	Weak (Delgado & van Schaik, 2000)	Weak (Robbins <i>et al.</i> , 2004; Hagemann <i>et al.</i> , 2018)	Strong (van Hooff & van Schaik, 1994; Watts & Mitani, 2001; Gilby & Wrangham, 2008; Mitani, 2009b; Surbeck <i>et al.</i> , 2017)	Weak (Kano, 1992; De Waal, 1995; Surbeck <i>et al.</i> , 2017)
	♀-♀	Weak (Medium if kin) (Delgado & van Schaik, 2000; Knott <i>et al.</i> , 2008; Singleton <i>et al.</i> , 2009)	Weak (Medium if kin) (Robbins <i>et al.</i> , 2004; Hagemann <i>et al.</i> , 2018)	Medium (Lehmann & Boesch, 2008, 2009; Langergraber <i>et al.</i> , 2013; Surbeck <i>et al.</i> , 2017)	Strong (Kano, 1992; De Waal, 1995; Surbeck <i>et al.</i> , 2017)
	♂-♀	Weak (Medium if kin) (Delgado & van Schaik, 2000; Knott <i>et al.</i> , 2008; Singleton <i>et al.</i> , 2009)	Strong (Robbins <i>et al.</i> , 2004; Hagemann <i>et al.</i> , 2018)	Medium (van Hooff & van Schaik, 1994; Watts & Mitani, 2001; Lehmann & Boesch, 2008; Gilby & Wrangham, 2008; Mitani, 2009b; Langergraber <i>et al.</i> , 2013; Surbeck <i>et al.</i> , 2017)	Strong (Kano, 1992; De Waal, 1995; Surbeck <i>et al.</i> , 2017)
<b>Inter-group competition</b>	Medium (♀) to Strong (♂)	Medium (Poole, 1987; Delgado & van Schaik, 2000; Knott <i>et al.</i> , 2008; van Noordwijk <i>et al.</i> , 2012)	Medium (Levréro, 2005; Harcourt & Stewart, 2007; Forcina <i>et al.</i> , 2019)	Strong (Wrangham, 1999; Mitani, 2009a; Mitani, Watts, & Amsler, 2010; Wilson <i>et al.</i> , 2012)	Weak (Kuroda, 1980; Furuichi, 2011)
<b>Percentage of lethal intra-specific aggressions</b>		0.08%	0.14%	4.49%	0.68%

		(Gómez <i>et al.</i> , 2016)	(Gómez <i>et al.</i> , 2016)	(Gómez <i>et al.</i> , 2016)	(Gómez <i>et al.</i> , 2016)	
	<b>Percentage of time spent in social grooming</b>	0% (Dunbar, 1991; McComb & Semple, 2005)	0.09% (Lehmann, Korstjens, & Dunbar, 2007)	11.7% (Lehmann <i>et al.</i> , 2007)	8.2% (Sakamaki, 2013)	
<b>Vocal behaviour</b>	<b>Vocal repertoire size (i.e. nb of call types)</b>	<b>Wild</b>	25 (Hardus <i>et al.</i> , 2009)	16 (Salmi <i>et al.</i> , 2013)	13 – 29 (Marler & Tenaza, 1977; Goodall, 1986; Arcadi, 2000; Slocombe & Zuberbühler, 2010)	15 (Bermejo & Omedes, 1999)
		<b>Captivity</b>	8 <sup>a</sup> <sub>OD</sub>	11 (Lemasson <i>et al.</i> , 2018)	10 <sup>a</sup> <sub>ZPB</sub>	12 (De Waal, 1988)
	<b>Hourly rate of vocal production (Group [G] / Individual [I] levels)</b>	<b>Wild</b>	Unstudied	G: 12.8 I: 8.7±5.4 (Salmi <i>et al.</i> , 2013)	G: 29.4 <sup>a</sup> <sub>BF</sub> I: 2.3±3.1 <sup>a</sup> <sub>BF</sub> - 2.6±2.0 (Arcadi, 2000)	G: 23.3 <sup>a</sup> <sub>NF</sub>
		<b>Captivity</b>	G: 1.9 <sup>a</sup> <sub>ZPB</sub> - 10.32 <sup>a</sup> <sub>OD</sub> I: 0.2±0.4 <sup>a</sup> <sub>ZPB</sub> - 1.3±1.4 <sup>a</sup> <sub>OD</sub>	G: 65 (Lemasson <i>et al.</i> , 2018)	G: 22.8 <sup>a</sup> <sub>ZPB</sub> - I: 1.32±1.7 <sup>a</sup> <sub>ZPB</sub>	G: 4.8 - 8.4 (De Waal, 1988; Levréro <i>et al.</i> , 2019)
<b>Calling patterns</b>	<b>Percentage of calls uttered in vocal interaction</b>	<b>Wild</b>	Unstudied	Unstudied	19% <sup>a</sup> <sub>BF</sub> (Florence/Clément, n.d.)	53 %
		<b>Captivity</b>	0% <sup>a</sup> <sub>ZPB/OD</sub>	38% (Lemasson <i>et al.</i> , 2018)	16% <sup>a</sup> <sub>ZPB</sub>	74% (Levréro <i>et al.</i> , 2019)
	<b>Main call type used during vocal interactions</b>	None <sup>a</sup> <sub>ZPB/OD</sub>	Grunts (turn-taking) (Harcourt & Stewart, 1996; Hedwig <i>et al.</i> , 2015a, 2015b)	Pant-hoots (overlap) <sup>a</sup> <sub>ZPB</sub> (Mitani & Nishida, 1993; Mitani & Brandt, 1994; Gros-Louis & Mitani, 1998; Fedurek <i>et al.</i> , 2013a; Fedurek, Schel, & Slocombe, 2013b; Fedurek, Donnellan, & Slocombe, 2014)	Hootings and barks (overlap) Peeps (turn-taking) (De Waal, 1988; Hohmann & Fruth, 1994; Bermejo & Omedes, 1999)	
	<b>Mean number of interacting partners</b>	0 <sup>a</sup> <sub>ZPB/OD</sub>	2.4±0.06 (Lemasson <i>et al.</i> , 2018)	0.34±0.7 <sup>a</sup> <sub>ZPB</sub>	2.2±0.5 (Levréro <i>et al.</i> , 2019)	

<b>Apart from vocal interactions</b>	<b>Percentage of isolated calls<sup>b</sup></b>	<b>Wild</b>	Unstudied	Unstudied	9% <sup>a-BF</sup>	24 % (Cornec <i>et al.</i> , submitted)
		<b>Captive</b>	14% <sup>a-ZPB</sup> 49% <sup>a-OD</sup>	9% (Lemasson <i>et al.</i> , 2018)	12% <sup>a-ZPB</sup>	5% (Levréro <i>et al.</i> , 2019)
	<b>Percentage of repeated calls<sup>c</sup></b>	<b>Wild</b>	Unstudied	Unstudied	72% <sup>a-BF</sup>	22 % (Cornec <i>et al.</i> , submitted)
		<b>Captive</b>	86% <sup>a-ZPB</sup> 51% <sup>a-OD</sup>	53% (Lemasson <i>et al.</i> , 2018)	70.5% <sup>a-ZPB</sup>	31% (Levréro <i>et al.</i> , 2019)
<b>Within vocal interactions</b>	<b>Percentage of calls in overlap<sup>d</sup></b>	<b>Wild</b>	Unstudied	Unstudied	10% <sup>a-BF</sup>	5% (Cornec <i>et al.</i> , submitted)
		<b>Captive</b>	0% <sup>a-ZPB/OD</sup>	10 % (Lemasson <i>et al.</i> , 2018)	11% <sup>a-ZPB</sup>	15% (Levréro <i>et al.</i> , 2019)
	<b>Organisation of overlap</b>	<b>Temporal</b>	None <sup>a-ZPB/OD</sup>	Unstudied	♂ adjust the constituent phases to promote chorusing (Fedurek <i>et al.</i> , 2013b)	Perfect synchronisation/alternation of <i>high-hoot</i> vocalisation (De Waal, 1988; Hohmann & Fruth, 1994)
		<b>Social</b>	None <sup>a-ZPB/OD</sup>	Unstudied	Interlocutors are mainly adults ♂ with affiliative bonds (Fedurek <i>et al.</i> , 2013a)	Interlocutors are both ♂ and ♀ with strong affiliative bonds (Hohmann & Fruth, 1994)
	<b>Percentage of calls in turn-taking<sup>e</sup></b>	<b>Wild</b>	Unstudied	Unstudied	9% <sup>a-BF</sup> (Arcadi, 2000)	50% (Cornec <i>et al.</i> , submitted)
		<b>Captive</b>	0% <sup>a-ZPB/OD</sup>	28% (Lemasson <i>et al.</i> , 2018)	5% of calls produced in all contexts <sup>a-ZPB</sup>	49% (Levréro <i>et al.</i> , 2019)
<b>Organisation of turn-taking</b>	<b>Temporal</b>	None <sup>a-ZPB/OD</sup>	Clear inter-calls delay of 3 s maximum (average 0.5 s) (Lemasson <i>et al.</i> , 2018)	Possible inter-calls delay of 1.5 s maximum (average 0.5 s) <sup>a-ZPB/BF</sup>	Clear inter-calls delay of 2.5 s maximum (average 0.4 s) (Levréro <i>et al.</i> , 2019)	

	Play-back experiment confirming the relevance of the turn-taking rule	Unstudied	Yes (Pougnault <i>et al.</i> , 2020b)	Unstudied	Yes (Levréro <i>et al.</i> , 2017)
	Social	None <sup>a</sup> <sub>ZPB/OD</sub>	High rate of interactions between the dominant adult ♂ and adult ♀, and low rate of interactions between ♀  (Lemasson <i>et al.</i> , 2018; Pougnault <i>et al.</i> , 2020b)	None <sup>a</sup> <sub>ZPB</sub> (Arcadi, 2000)	Adult of both sexes (66%) and sub-adults ♂ (37%) (Cornec <i>et al.</i> , submitted)  Interlocutors with strong affiliative social bonds (Levréro <i>et al.</i> , 2019)

**Unstudied:** subject not addressed in the current literature.

**None:** absence of this type of vocal pattern in the studied species according to the current literature or our own data set.

<sup>a</sup> Calling patterns from data of wild chimpanzees from the Budongo Forest Reserve, Uganda (<sup>a</sup><sub>BFR</sub>), wild bonobos from Nkala forest in Bolobo Territory, Democratic Republic of Congo (<sup>a</sup><sub>NF</sub>) and captive chimpanzees and orangutans from ZooParc de Beauval, France (<sup>a</sup><sub>ZPB</sub>) and Ouwehand Direnparc, Netherlands (<sup>a</sup><sub>OD</sub>). The details of these unpublished data are given in *Supporting information 1, 2 and 3*.

<sup>b</sup> Isolated call occurs when one call is emitted and no other calls can be heard around, according to Bouchet *et al.*, 2010.

<sup>c</sup> Repeated calls occur when the same caller calls several times in a row, according to Koda (2004).

<sup>d</sup> Overlapped calls correspond to a partial (or total) superposition of calls from two or more callers.

<sup>e</sup> Turn-taking-based vocal interactions correspond to an orderly exchange of communicative signals between individuals adjusting the timing of their vocal responses by respecting a minimum silence gap (preventing call overlap), and a maximum silence gap (allowing to ensure call coordination).

Symbols: ♂ = male; ♀ = female.

## 2.1. Origins and geographical repartition

Genetic evidence indicates that the last common ancestor to *Ponginae* and *Homininae* lived about 9 to 18 million years ago. A second speciation, between ~8.6 and 11.5Myr, led to the apparition of *Gorillini* and *Hominini*, and a final differentiation, between ~4.2 and 9Myr, gave rise to *Paninina* and *Hominina* clades. Given the phylogenetic proximity and cladistics pedigree, the *Pan* genus thus represents our closest living relatives.

Despite a growing body of evidence defending the idea of an European origin of great apes (e.g. Begun, 2010; Böhme *et al.*, 2019), all currently living non-human great apes are located in tropical and equatorial rain forests of Asia and Africa. Some of them are more tolerant regarding their habitat types. Bornean orang-utans live almost exclusively in the canopy of humid forests, while Western lowland gorillas are mainly terrestrial and are encountered in both humid and *terra firma* lowland forests (see Table 1). Chimpanzees and bonobos mainly occupy primary and secondary humid forests but can sometimes be found on savannah woodland mosaics. They evolve in the trees, as well as on the ground according to their age. As a result, all study species live most of the time in habitats with reduced visibility (due to the density of the surrounding vegetation) and travel in a more or less 3-dimensional space. These characteristics are known to be particularly suitable for vocal communication as acoustic signals can be easily transmitted to neighbouring partners (Bradbury & Vehrencamp, 2011; Lemasson, 2011).

## 2.2. Social systems

The four great apes live in four different social systems and all differ in group size, group structure, and social organization.

Bornean orang-utans are known to be the most solitary living great ape species. Males live generally alone and females can typically be seen along with their infants. Despite some

non-random temporary peaceful grouping observed near fruit sources (van Schaik, Marshall, & Wich, 2009), a high level of intolerance is generally reported between males notably for access to females (Galdikas, 1985). All three other species form clear social groups.

In contrast to orang-utans, which are female-philopatric, chimpanzees and bonobos are male-philopatric species living in large multi-male multi-female and highly flexible fission-fusion societies (i.e. consist of subgroups of variable size and composition in which group members regularly join –fusion– or separate from –fission– one another, Kummer, 1971). Subgroups are typically composed of 20 to 100 individuals. Both chimps and bonobos have a high rate of physical interactions, including at least eight times higher grooming rates than the two other great apes. Male chimpanzees are highly territorial. Related males form strong social bonds on which they rely to manage important inter- and intra-community competition, while females remain more in the periphery and have weak social bonds with each other. This high level of competition leads to aggressive conflicts in chimpanzees and the highest (at least four times more) recorded death rate due to intra-individual conflicts among great apes. Also, chimpanzees appear to be less prone to cooperative behaviours than their bonobo relatives (Hare & Tomasello, 2004; Jensen *et al.*, 2006; Melis, Hare, & Tomasello, 2006; Hare *et al.*, 2007; Hirata & Fuwa, 2007; Pelé *et al.*, 2009; Gruber & Clay, 2016; Suchak *et al.*, 2016), for whom inter- and intragroup interactions are mainly tolerant. In bonobos, members from different communities regularly mix peacefully to feed, and females have a central role in the community. They establish strong affiliative bonds with other females independently of their relative kinship and also form bonds with unrelated males.

Western lowland gorillas live in uni-male/multi-female groups of middle size (i.e. 10 to 20 individuals) where both sexes disperse. The sole adult male develops strong social bonds with adult females. He is the key leader of the group, coordinating activities, protecting the group and managing social interactions between females (Parnell, 2002; Harcourt & Stewart,

2007). Despite the strong social bonds that unite male and female gorillas, individuals appear to be poorly cooperative (Harcourt & Stewart, 2007; Pelé *et al.*, 2009).

Living in different societies obviously triggers different social needs in term of daily social interactions and it is important to consider how this may influence vocal interaction patterns. Interestingly, all great ape species are known to have high and comparable cognitive skills (even social) (Chalmeau *et al.*, 1997; Pelé *et al.*, 2009; Dufour *et al.*, 2009), suggesting that if different vocal rules exist in the four species, it has certainly more to do with social differences than with habitat and cognitive differences.

### 3. Vocal interactions in great apes

#### 3.1. Repertoire size and hourly rate of vocal production

Following the “social complexity hypothesis for communication”, we would expect to see an increasing number of distinct call types in the vocal repertoires of social species, increasing from orang-utans to gorillas and then to chimpanzees and bonobos if we refer to group sizes and time spent grooming (e.g. McComb & Semple, 2005; Gustison *et al.*, 2012). But this is clearly not the case. Wild-based studies showed that the repertoire size of Bornean orang-utans and chimpanzees is almost similar (respectively 25 for orang-utans and 13 to 29 for chimpanzees according to authors and studied populations mentioned in the Table1). Moreover, a comparable vocal repertoire size is found in wild bonobos and gorillas with respectively 15 and 16 different call types. In captivity, the repertoire size of each great ape species contains less distinct call types than in wild populations. All species show a reasonably equivalent number of distinct call types in captivity (*i.e.* between 8 and 12).

The hourly rate of vocal production remains high for all great ape species but strong variations can be found between study populations and across environmental conditions. The environmental stimulations encountered in the wild are indeed much more diverse and frequent



than in captivity which typically results in more call types and higher call rates in free-ranging populations (Lameira & Call, 2020). As for vocal repertoire size, the link between hourly vocal rate and social complexity appears very weak. This questions the relevance of restricting social complexity to group size and grooming rate measures in the debated theories (McComb & Semple, 2005; Freeberg *et al.*, 2012; Kershenbaum, Freeberg, & Gammon, 2015). This is particularly striking when considering the very large repertoire size of orang-utans (Hardus *et al.*, 2009). It is of course possible that orang-utans were more social in the distant past than they are at present (Rijksen, 1978).

Some methodological issues, referring to the classification of call types, also have to be addressed. Many communicative repertoires reveal intermediate forms between specific signals as well as variations within signal types. These so-called graded repertoires are characteristic of great apes, particularly chimpanzees and bonobos, leading to difficulties in identifying a clear-cut number of stereotypic signals in a species repertoire (Keenan *et al.*, 2020). Experimental approaches using playback and systematic contextual studies of call emission are still needed to test whether simple acoustic analyses induce a possible over or underestimation of repertoire size.

The question about whether vocal interaction pattern measures are better suitable candidates to support the “social complexity hypothesis for communication” thus remains open at this stage, and this is the aim of the following section.

### 3.2. Occurrences and features of calling patterns across great ape species

The three calling patterns previously introduced are not randomly distributed across the four great ape study species. Vocal interactions are frequent in bonobos (more than 50 % of their vocal production occurs in calm contexts), followed by gorillas (around 40%) and then chimpanzees (less than 20%), and they seem inexistent in orang-utans. Vocal interactions,

regardless of the species average group size, typically concern only two (sometimes three) coordinating partners. This largely parallels what is found in monkeys' conversation-like vocal exchanges (Pougnault *et al.*, 2020a). The typical pattern of orang-utan vocal production is isolated calling. Chimpanzees also often emit isolated calls, but most of their vocal productions consist of series of repeated calls.

When focusing on the organisation of vocal interactions, differences can also be found in the three species who interact vocally. Chimpanzees differ from the two others: overlapping seems to be their rule of interacting (about 2/3 of their vocal interactions are overlapped), particularly for males. In a former study, Arcadi (2000) also concluded in favour of an intriguing quasi-inexistence of turn-taking in free-ranging chimpanzees. Chimpanzees are indeed better known to form structured choruses, mainly based on a superposition of *pant-hoots* (Gros-Louis & Mitani, 1998; Fedurek *et al.*, 2013b; Ravignani *et al.*, 2014), which seem to play a key role in territory defence and in regulating grouping dynamics facilitating fusion and changes in sub-group composition (Fedurek *et al.*, 2014). A pant-hoot vocalization is composed of four distinct phases (Gros-Louis & Mitani, 1998; Slocombe & Zuberbühler, 2010): (i) an introduction consisting of low frequency and low amplitude elements, (ii) a build-up phase comprising a series of short inhalations and exhalations, (iii) a climax which correspond to the loudest part of the call including one or several screams in succession, and sometimes (iv) a short let-down phase that has similar features as the build-up. During pant-hoot chorusing, interlocutors adjust the duration of constituent phases (i.e. build-up and climax) promoting choruses in a way to mirror the partner's call, revealing that the temporal structure of this rather stereotyped vocalisation is sensitive to fine details of the vocal behaviour of the audience (Fedurek *et al.*, 2013b).

Conversely, bonobos, and to lesser extent gorillas, tend to avoid overlap when they interact vocally (about 66% and up to 94% of the vocal interactions without overlap in bonobos

in captivity and in the wild, respectively, and 74% in gorillas, Figure 1). As it was found in monkeys previously, gorillas and bonobos engage frequently in vocal exchanges respecting a turn-taking rule. Their vocal exchanges are call type-dependent. For both species, soft contact calls, known to be characteristic of peaceful contexts (*i.e.* gorillas: grunts/double grunt; bonobos: peeps/peep-yelps), are predominantly used. The inter-call duration between two consecutive interlocutors is also species-specific. It is of 3 seconds maximum in gorillas (0.5 s optimum Lemasson *et al.*, 2018) and between 2.1 and 2.5 seconds maximum in bonobos (wild bonobos: 0.7 s optimum, Cornec *et al.*, submitted, and captive bonobos: 0.4 s optimum, Levréro *et al.*, 2019 respectively). This inter-call duration is consistent with those found in several monkey species, *i.e.* rarely shorter than the mean duration of a call, although most published maximum response delays are shorter in monkeys (around 1 or 1.5 s). Recent playback experiments, based on the paradigm of violation-of-expectations, conducted on gorillas and bonobos revealed that individuals perceived when the short silence gap between vocalisations produced by different interlocutors is present or not (Levréro *et al.*, 2017; Pournault *et al.*, 2020b). These results suggest that the respect of response delay matters to the audience (Levréro *et al.*, 2017; Pournault *et al.*, 2020b).

Our analysis does not support a strict correlation between group size, grooming rate and vocal exchange rate, as it could be expected from Dunbar's "grooming-at-distance" hypothesis (1996). However, a subtler link between the nature of the society and the type of vocal interaction emerges from our comparative data. Call overlap and turn-taking exchanges appear as two possible vocal alliance strategies to advertise or strengthen social bonds as detailed below.

### 3.3. Social features of vocal interactions

Regarding chimpanzee choruses or gorilla and bonobo call exchanges, we see that interlocutors are not randomly chosen.

In chimpanzees, pant-hoot choruses are mainly produced by high-ranking males (Mitani & Nishida, 1993; Clark & Wrangham, 1994), with the aim of recruiting and maintaining the company of close associates (Mitani & Nishida, 1993). Researchers highlight that choruses occur between allies without consideration of kinship, or age (Mitani & Nishida, 1993). Indeed, this vocal behaviour appears as a strong indicator of affiliation between males (Fedurek *et al.*, 2013a), and during the participation of a chorus, allies tend to match vocally the acoustic features of a partner's calls (Mitani & Brandt, 1994; Gros-Louis & Mitani, 1998). Then, pant-hoot choruses might be a good low-cost behaviour in highly fluid fission-fusion societies (Fedurek *et al.*, 2013a). Also, the prevalence of *pant-hoot* chorusing in chimpanzee vocal interactions could be based on the fact that, in a context of highly risky intra- and inter-community competition, opponents seem able to perceive the number of participants of a *pant-hoot* chorus and to adapt their behaviour in response to this detection (Wilson, Hauser, & Wrangham, 2001). Attesting social affinities, the loud calls produced in chorus might deter potential rivals and avoid a costly fight (Wilson *et al.*, 2001).

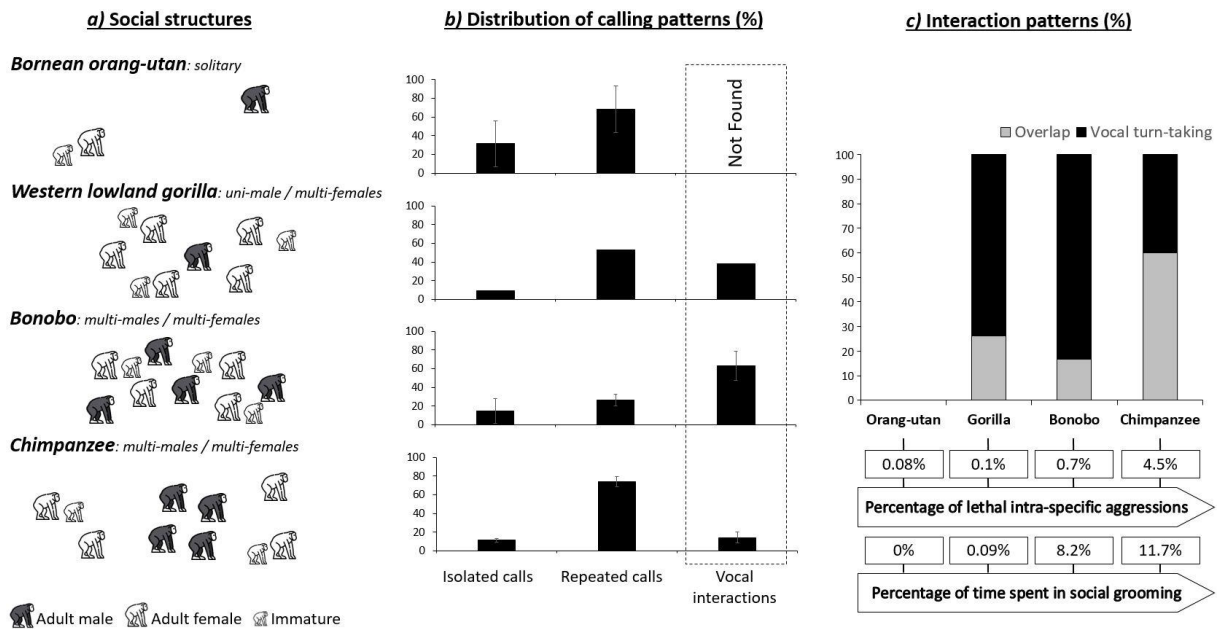
Vocal exchanges in gorillas and bonobos are also socially-guided. In Western lowland gorillas, the adult male vocally responds preferentially to adult females with whom he establishes durable social bonds, and less with individuals distant in age such as sub-adults or juveniles (Lemasson, Pereira, & Levréro, 2018). The same is true for bonobos for which vocal turn-taking occurs mainly between interlocutors with strong social affinities, and is not influenced by the age, the sex, the hierarchical rank or the kinship of interlocutors (Levréro *et al.*, 2019). Experimental approaches, using violation-of-expectation paradigms, also tested the relevance of social rules previously described in bonobos and Western lowland gorillas.

Authors compared the attentional response of the audience towards vocal exchanges implying individuals with or without strong social affinities in bonobos (Levréro *et al.*, 2017) and close or distant in age in gorillas (Pougnault *et al.*, 2020a). It seems clear that individuals have some expectations about which individuals are supposed to vocally interact (Pougnault *et al.*, 2020a).

So, our focused phylogenetic comparisons largely support the social complexity hypothesis for communication as suggested by several others authors (e.g. Blumstein & Armitage, 1997; Freeberg *et al.*, 2012; Pollard & Blumstein, 2012; Manser *et al.*, 2014), but only by considering a more complex set of subtle features than what is generally brought to light. Indeed, if we focus our attention only on the number of different call types composing the vocal repertoire of each great ape species with regards to the nature of social interactions or the number of group members, the results obtained are not convincing and the methodology can be criticized on various aspects (e.g. lack of consideration of relationship types between group members, or the subjective classification of calls from a graded repertoire: Fischer *et al.*, 2017a, 2017b; Hammerschmidt & Fischer, 2019). However, it seems possible to structure a gradation in social-vocal complexity considering the three associative vocal-social patterns described following this investigation (see Figure 1). Indeed, the Bornean orang-utans are mainly solitary and favour the production of isolated calls. In chimpanzee societies, which are mainly despotic with a strong influence of males, consecutive calls from single callers (i.e. characteristic of emotional discharge: Lemasson *et al.*, 2010b; Owren & Rendall, 1997; Scherer, 2003) are particularly common, and vocal interactions occur frequently in the form of choruses of pant-hoots. On the contrary, gorillas and bonobos, who are more social than orang-utans and have more peaceful interactions with their group members than chimpanzees, present the majority of vocal interactions characterized by overlap avoidance and turn-taking. Interestingly, chimpanzees spend more time in social grooming than the other three species (see Figure 1). They appear to use dyadic mutual grooming (Fedurek & Dunbar, 2009) in a fluid way to access

a large number of grooming partners, by changing partners more often than do bonobos (Girard-Buttoz *et al.*, 2020). Moreover, a large part of their time spent in social grooming is also allotted to polyadic grooming (16% of the total grooming time for chimpanzees (Nakamura, 2003) against only 3% in bonobos (Sakamaki, 2013)). This polyadic grooming involves a larger number of partners in chimpanzees than in bonobos (Girard-Buttoz *et al.*, 2020). By doing so, chimpanzees reach the same efficiency as humans' conversations in terms of number of partners accessed simultaneously in polyadic grooming clusters (i.e. 2.72 partners in humans and 2.18 partners in chimpanzees; Nakamura, 2000). Paradoxically, these clusters occur particularly between males of similar hierarchical rank in chimpanzees, while bonobos engage more often with conspecifics more distant in dominance rank (Girard-Buttoz *et al.*, 2020).

At a first glance, the low rate of vocal exchanges in chimpanzees appears counter-intuitive given the high rate of social grooming (see Figure 1). But, the apparent efficiency of polyadic social grooming, the limited number of close social partners coupled with the high level of intra-group competition could explain why this species do not need to “groom-at-distance” using vocal exchanges to attest their strong relationship (Dunbar, 1996). In contrast, in species displaying more tolerance among group members and less time spent in social grooming, as bonobos and gorillas, vocal exchanges that are temporally and socially ruled are favoured during vocal interactions. Nonetheless, this interpretation must be taken with caution given the weak representation of each social system here, and highlights the need for a multi-factorial approach to social parameters.



**Figure 1** | Synthetic representation of social structures, distribution of calling patterns and interaction patterns according to the percentage of lethal intra-specific aggressions and the time spent in social grooming for each of the four studied species. The data used to build the graphs of panels (b) and (c), respectively untitled “Distribution of calling patterns (%)” and “Interactions patterns (%)” were obtained by average of data collected from wild and captivity when it is possible. All the references used to prepare the figure presented here can be found in Table 1.

#### 4. Conclusion

Our review suggests that vocal interaction patterns match the quality of social bonds in apes, thus supporting theories of a socially-driven evolution. A well-targeted comparative approach can help to identify the multiple and subtle social factors (beyond only considering the social structure or organization) underlying social complexity that drive the evolution of vocal communication. While focusing on great apes gives access to only one template for each different social life intensity, we believe that focusing on our closest living relatives is one

necessary step that may help to better understand the social factors influencing different vocal interaction patterns.

It remains difficult to draw an evolutionary scenario of vocal interaction skills, from the *Strepsirhini* to *Homo* genus, because the emergence of social and vocal complexity is profuse in the primate lineage and probably results from multiple evolutive pressures. However, great apes and monkeys share some temporal and social rules regulating their vocal exchanges, which may show similar constraints between primate “conversational” vocal exchanges and human conversations. It might be possible that chimpanzees and orang-utans have been subject to different selective pressures, respectively strong male despotic hierarchy and a mainly solitary life, leading to the appearance of derived evolutionary traits.

Given the large body of literature on the use of gestures and facial expression in chimpanzees (e.g. Graham *et al.*, 2018; Roberts & Roberts, 2019; vanLawick-Goodall, 1968), but also in other apes (e.g. orang-utans Liebal, Pika, & Tomasello, 2006; gorillas Genty *et al.*, 2009; bonobos Graham, Furuichi, & Byrne, 2017; bonobos and chimpanzees Graham *et al.*, 2018), and its importance in human communication (Goldin-Meadow, 1999; McNeill, 2000; Kendon, 2004; Holle & Gunter, 2007), it seems crucial for further studies to develop more multimodal approaches to understand the communicative interactions of apes. In this way, a promising study conducted by Fröhlich and colleagues (2016) on *Pan* species highlighted the ability of chimpanzees to engage in turn-taking not involving vocalisations. Thus, conclusions about turn-taking skills cannot be drawn from the sole presence or absence of vocal turn-taking interactions.

Here, a vast research area is opening into the investigation of the function of these different vocal interaction strategies. The nature of messages contained in vocal interactions - if there is exchanged information apart from the caller identity or its spatial position- still need to be explored. For instance, experimental approaches using playback tests should be conducted



more systematically. Yet, it is clear that the ability to interact vocally with others, even limited to short vocal exchanges as we have shown, seems deeply rooted in the animal lineage and could have been an important step towards the complexification of social life, and the evolution of animal vocal communication (Dunbar, 1996).

## Acknowledgments

This work was financially supported by a CIFRE grant (#2018\_00074) funded by the Association Beauval Nature pour la Conservation et la Recherche, Centre National de la Recherche Scientifique, Rennes 1 University and Lyon/Saint-Etienne University. This research was also financially supported by the University of Lyon/Saint-Etienne, *Nature & Découverte* foundation and WWF-Belgium. We are deeply grateful to the ZooParc de Beauval, the Ouwehand Dierenpark and the orang-utans caregivers for their valuable assistance. Invaluable help in the field was provided by the field assistants of WWF-DRC, of the NGO Mbou-Mon-Tour and the management and staff of the Budongo Conservation Field Station. We also warmly thank Jean-Christophe Bokika, president of the NGO Mbou-Mon-Tour, Raymond Lumbuenamo, national coordinator of the WWF-DRC and Damien Vincent, general manager of the WWF-Belgium and the chief of the village of Nkala for their support and permission to study the bonobos in the Nkala forest. We also thank the local community at Nkala for the support which was greatly appreciated. We are very grateful to the Uganda Wildlife Authority and the Uganda National Council for Science and Technology for permission to conduct the study. We thank Katarzyna Pisanski for correcting our English.

## References

- ANCRENAZ, M., GUMAL, M., MARSHALL, A.J., MEIJAARD, E., WICH, S.A. & HUSSON, S. (2018) *Pongo pygmaeus* (errata version published in 2018). The IUCN Red List of Threatened Species 2016.
- ARCADI, A.C. (2000) Vocal responsiveness in male wild chimpanzees: Implications for the evolution of language. *Journal of Human Evolution* **39**, 205–223.
- ARLET, M., JUBIN, R., MASATAKA, N. & LEMASSON, A. (2015) Grooming-at-a-distance by exchanging calls in non-human primates. *Biology Letters* **11**, 20150711–20150714.
- ARORA, N., NATER, A., VAN SCHAIK, C.P., WILLEMS, E.P., VAN NOORDWIJK, M.A., GOOSSENS, B., MORF, N., BASTIAN, M., KNOTT, C., MORROGH-BERNARD, H., KUZE, N., KANAMORI, T., PAMUNGKAS, J., PERWITASARI-FARAJALLAH, D., VERSCHOOR, E., ET AL. (2010) Effects of Pleistocene glaciations and rivers on the population structure of Bornean orangutans (*Pongo pygmaeus*). *Proceedings of the National Academy of Sciences of the United States of America* **107**, 21376–21381.
- BEGUN, D.R. (2010) Miocene Hominids and the Origins of the African Apes and Humans. *Annual Review of Anthropology* **39**, 67–84. Annual Reviews.
- BERMEJO, M. & OMEDES, A. (1999) Preliminary Vocal Repertoire and Vocal Communication of Wild Bonobos (*Pan paniscus*) at Lilungu (Democratic Republic of Congo). *Folia Primatologica* **70**, 328–357.

- BIBEN, M., SYMMES, D. & MASATAKA, N. (1986) Temporal and Structural Analysis of Affiliative Vocal Exchanges in Squirrel Monkeys ( *Saimiri sciureus* ). *Behaviour* **98**, 259–273.
- BLUMSTEIN, D.T. & ARMITAGE, K.B. (1997) Does sociality drive the evolution of communicative complexity? A comparative test with ground-dwelling sciurid alarm calls. *The American Naturalist* **150**, 179–200.
- BÖHME, M., SPASSOV, N., FUSS, J., TRÖSCHER, A., DEANE, A.S., PRIETO, J., KIRSCHER, U., LECHNER, T. & BEGUN, D.R. (2019) A new Miocene ape and locomotion in the ancestor of great apes and humans. *Nature* **575**, 489–493. Nature Research.
- BOUCHET, H., BLOIS-HEULIN, C. & LEMASSON, A. (2013) Social complexity parallels vocal complexity: A comparison of three non-human primate species. *Frontiers in Psychology* **4**, 1–15.
- BOUCHET, H., KODA, H. & LEMASSON, A. (2017) Age-dependent change in attention paid to vocal exchange rules in Japanese macaques. *Animal Behaviour* **129**, 81–92. Academic Press.
- BOUCHET, H., PELLIER, A.S., BLOIS-HEULIN, C. & LEMASSON, A. (2010) Sex differences in the vocal repertoire of adult red-capped mangabeys (*Cercocebus torquatus*): A multi-level acoustic analysis. *American Journal of Primatology* **72**, 360–375.
- BRADBURY, J.W. & VEHCAMP, S.L. (2011) *Principles of animal communication (2nd ed.)*. Sinauer Associates.
- BRISÉÑO-JARAMILLO, M., RAMOS-FERNÁNDEZ, G., PALACIOS-ROMO, T.M., SOSA-LÓPEZ, & J.R., LEMASSON, & A., SOSA-LÓPEZ, J.R. & LEMASSON, A. (2018) Age and social affinity effects on contact call interactions in free-ranging spider monkeys. *Behavioral Ecology and Sociobiology* **72**, 192.
- BRISÉÑO JARAMILLO, M., GUYOT-BIQUAND, V., ESTRADA, A. & LEMASSON, A. (2017) Vocal repertoire of free-ranging black howler monkeys' (*Alouatta pigra*): call types, contexts and sex-related contributions. *American Journal of Primatology* **79**, 22630–10. Wiley.
- CANDIOTTI, A., ZUBERBUHLER, K. & LEMASSON, A. (2012) Convergence and divergence in Diana monkey vocalizations. *Biology Letters* **8**, 382–385.
- CARTER, G.G., SKOWRONSKI, M.D., FAURE, P.A. & FENTON, B. (2008) Antiphonal calling allows individual discrimination in white-winged vampire bats. *Animal Behaviour* **76**, 1343–1355.
- CASTELLANO, D. & MUNCH, K. (2020) Populations Genomics in the Great apes. In *Statistical Population Genomics* (ed JULIEN Y. DUTHEIL), pp. 453–468 Humana Press. Springer Protocols, Plön, Germany.
- CHALMEAU, R., LARDEU, K., BRANDILLAS, P. & GALLO, A. (1997) Cooperative Problem Solving by Orangutans (*Pongo pygmaeus*). *International Journal of Primatology* **18**, 23–32.
- CHEN, H.C., KAPLAN, G. & ROGERS, L.J. (2009) Contact calls of common marmosets (*Callithrix jacchus*): Influence of age of caller on antiphonal calling and other vocal responses. *American Journal of Primatology* **71**, 165–170.
- CHENEY, D.L., SEYFARTH, R.M. & SILK, J.B. (1995) The Responses of Female Baboons (*Papio cynocephalus ursinus*) to Anomalous Social Interactions: Evidence for Causal Reasoning? *Journal of Comparative Psychology* **109**, 134–141.
- CHOW, C.P., MITCHELL, J.F. & MILLER, C.T. (2015) Vocal turn-taking in a non-human primate is learned during ontogeny. *Proceedings of the Royal Society B: Biological Sciences* **282**, 20150069.
- CLARK, A.P. & WRANGHAM, R.W. (1994) Chimpanzee Arrival Pant-Hoots: Do Signify Food or Status? *International Journal of Primatology* **15**, 185–205.

- DARWIN, C. (1877) L'expression des émotions chez l'homme et les animaux. *Gallica, Bibliothèque Nationale de France*, 203.
- DELGADO, R.A. & VAN SCHAIK, C.P. (2000) The behavioral ecology and conservation of the orangutan (*Pongo pygmaeus*): A tale of two islands. *Evolutionary Anthropology: Issues, News, and Reviews* **9**, 1520–6505.
- DROSOPOULOS, S. & CLARIDGE, M.F. (2005) *Insect Sounds and COmmunication: Physiology, Behaviour, Ecology and Evolution*. CRC Press, London.
- DUFOUR, V., PELÉ, M., NEUMANN, M., THIERRY, B. & CALL, J. (2009) Calculated reciprocity after all: Computation behind token transfers in orang-utans. *Biology Letters* **5**, 172–175. Royal Society.
- DUNBAR, R.I.M. (1991) Functional significance of social grooming in primates. *Folia Primatologica* **57**, 121–131.
- DUNBAR, R.I.M. (1996) *Grooming, gossip and the evolution of language*. Faber & Faber, London, United Kingdom.
- DUNBAR, R.I.M. (1998) The Social Brain Hypothesis. *Evolutionary Anthropology* **6**, 178\_190.
- VAN ELSACKER, L., VERVAECKE, H. & VERHEYEN, R.F. (1995) A Review of Terminology on Aggregation Patterns in Bonobos (*Pan paniscus*). In *International Journal of Primatology* p.
- FEDUREK, P., DONNELLAN, E. & SLOCOMBE, K.E. (2014) Social and ecological correlates of long-distance pant hoot calls in male chimpanzees. *Behavioral Ecology and Sociobiology* **68**, 1345–1355. Springer Verlag.
- FEDUREK, P. & DUNBAR, R.I.M. (2009) What does mutual grooming tell us about Why chimpanzees groom? *Ethology* **115**, 566–575.
- FEDUREK, P., MACHANDA, Z.P., SCHEL, A.M. & SLOCOMBE, K.E. (2013a) Pant hoot chorusing and social bonds in male chimpanzees. *Animal Behaviour* **86**, 189–196. Elsevier Ltd.
- FEDUREK, P., SCHEL, A.M. & SLOCOMBE, K.E. (2013b) The acoustic structure of chimpanzee pant-hooting facilitates chorusing. *Behavioral Ecology and Sociobiology* **67**, 1781–1789.
- FISCHER, J., FARNWORTH, M.S., SENNHENN-REULEN, H. & HAMMERSCHMIDT, K. (2017a) Quantifying social complexity. *Animal Behaviour* **130**, 57–66. Academic Press.
- FISCHER, J., WADEWITZ, P. & HAMMERSCHMIDT, K. (2017b) Structural variability and communicative complexity in acoustic communication. *Animal Behaviour* **134**, 229–237. Academic Press.
- FLORENCE/CLÉMENT (undated) Data Clément.
- FORCINA, G., VALLET, D., LE GOUAR, P.J., BERNARDO-MADRID, R., ILLERA, G., MOLINA-VACAS, G., DRÉANO, S., REVILLA, E., RODRÍGUEZ-TEIJEIRO, J.D., MÉNARD, N., BERMEJO, M. & VILÀ, C. (2019) From groups to communities in western lowland gorillas. *Proceedings of the Royal Society B: Biological Sciences* **286**, n/a.
- FREEBERG, T.M., DUNBAR, R.I.M. & ORD, T.J. (2012) Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 1785–1801.
- FREEBERG, T.M. & KRAMS, I. (2015) Does social complexity link vocal complexity and cooperation? Springer Verlag. *Journal of Ornithology* . .
- FRÖHLICH, M., KUCHENBUCH, P., MÜLLER, G., FRUTH, B., FURUICHI, T., WITTIG, R.M. & PIKA, S. (2016) Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-

- taking sequences. *Scientific Reports* **6**, 1–14.
- FRUTH, B., HICKEY, J.R., ANDRÉ, C., FURUICHI, T., HART, J., HART, T., KUEHL, H., MAISELS, F., NACKONEY, J., REINARTZ, G., SOP, T., THOMPSON, J. & WILLIAMSON, E.A. (2016) Pan paniscus (errata version published in 2016). The IUCN Red List of Threatened Species 2016.
- FRUTH, B., WILLIAMSON, E.A. & RICHARDSON, M.C. (2013) Bonobo Pan paniscus. In *Handbook of the Mammals of the World Volume 3: Primates* (ed A.B.R. AND D.E.W. R.A. MITTERMEIER), pp. 853–854 Lynx Edicions. Barcelona, Spain.
- FURUICHI, T. (2011) Female contributions to the peaceful nature of bonobo society. *Evolutionary Anthropology* **20**, 131–142.
- FURUICHI, T., IDANI, G., IHOBE, H., HASHIMOTO, C., TASHIRO, Y., SAKAMAKI, T., MULAVWA, M.N., YANGOZENE, K. & KURODA, S. (2012) Long-term studies on wild bonobos at Wamba, Luo Scientific Reserve, D. R. Congo: Towards the understanding of female life history in a male-philopatric species. In *Long-Term Field Studies of Primates* (ed HEIDELBERG), pp. 413–433 Springer. Springer-Verlag Berlin Heidelberg, Berlin.
- GALDIKAS, B.M.F. (1985) Adult male sociality and reproductive tactics among orangutans at Tanjung Puting. *Folia Primatologica* **45**, 9–24.
- GARCIA, M., THEUNISSEN, F., SÈBE, F., CLAVEL, J., RAVIGNANI, A., MARIN-CUDRAZ, T., FUCHS, J. & MATHEVON, N. (2020) Evolution of communication signals and information during species radiation. *Nature Communications* **11**, 4970.
- GATTI, S., LEVRÉRO, F., MÉNARD, N. & GAUTIERHION, A. (2004) Population and group structure of western Lowland Gorillas (*Gorilla gorilla gorilla*) at Lokoué, Republic of Congo. *American Journal of Primatology* **63**, 111–123.
- GEISSMANN, T. (1999) Duet songs of the siamang, *Hylobates syndactylus*: II. Testing the pair-bonding hypothesis during a partner exchange. *Behaviour* **136**, 1005–1039.
- GEISSMANN, T. & ORGELDINGER, M. (2000) The relationship between duet songs and pair bonds in siamangs, *Hylobates syndactylus*. *Animal Behaviour* **60**, 805–809.
- GENTY, E., BREUER, T., HOBAITER, C. & BYRNE, R.W. (2009) Gestural communication of the gorilla (*Gorilla gorilla*): Repertoire, intentionality and possible origins. *Animal Cognition* **12**, 527–546.
- GILBY, I.C. & WRANGHAM, R.W. (2008) Association patterns among wild chimpanzees (*Pan troglodytes schweinfurthii*) reflect sex differences in cooperation. *Behavioral Ecology and Sociobiology* **62**, 1831–1842.
- GINSBURG, G.P. & KILBOURNE, B.K. (1988) Emergence of vocal alternation in mother-infant interchanges. *Journal of Child Language* **15**, 221–235.
- GIRARD-BUTTOZ, C., SURBECK, M., SAMUNI, L., BOESCH, C., FRUTH, B., CROCKFORD, C., HOHMANN, G. & WITTIG, R.M. (2020) Variable use of polyadic grooming and its effect on access to social partners in wild chimpanzees and bonobos. *Animal Behaviour* **168**, 211–224. Academic Press.
- GOLDIN-MEADOW, S. (1999) The role of gesture in communication and thinking. *Trends in Cognitive Sciences* **3**, 419–429.
- GÓMEZ, J.M., VERDÚ, M., GONZÁLEZ-MEGÍAS, A. & MÉNDEZ, M. (2016) The phylogenetic roots of human lethal violence. *Nature* **538**, 233–237. Nature Publishing Group.
- GOODALL, J. (1986) *The Chimpanzees of Gombe: Patterns of Behavior*. Belknap Press of Harvard University Press, Cambridge.
- GRAFE, T.U. (1996) The function of call alternation in the African reed frog (*Hyperolius marmoratus*):

- precise call timing prevents auditory masking. *Behavioral Ecology and Sociobiology* **38**, 149–158.
- GRAFE, T.U. (1999) A function of synchronous chorusing and a novel female preference shift in an anuran. *Proceedings of the Royal Society B* **266**, 2331–2336.
- GRAHAM, K.E., FURUICHI, T. & BYRNE, R.W. (2017) The gestural repertoire of the wild bonobo (*Pan paniscus*): a mutually understood communication system. *Animal Cognition* **20**, 171–177. Springer Verlag.
- GRAHAM, K.E., HOBAITER, C., OUNSLEY, J., FURUICHI, T. & BYRNE, R.W. (2018) Bonobo and chimpanzee gestures overlap extensively in meaning. *PLoS Biology* **16**. Public Library of Science.
- GREENFIELD, M.D. (1994) Cooperation and Conflict in the Evolution of Signal Interactions. *Annual Review of Ecology, Evolution, and Systematics* **25**, 97–126.
- GROS-LOUIS, J. & MITANI, J. (1998) Chorusing and call convergence in chimpanzees: Tests of three hypotheses. *Behaviour* **135**, 1041–1064.
- GRUBER, T. & CLAY, Z. (2016) A Comparison Between Bonobos and Chimpanzees: A Review and Update. *Evolutionary Anthropology* **25**, 239–252.
- GUSTISON, M.L., LE ROUX, A. & BERGMAN, T.J. (2012) Derived vocalizations of geladas (*Theropithecus gelada*) and the evolution of vocal complexity in primates. *Philosophical Transactions of the Royal Society B: Biological Sciences* **367**, 1847–1859. Royal Society.
- HAGEMANN, L., BOESCH, C., ROBBINS, M.M., ARANDJELOVIC, M., DESCHNER, T., LEWIS, M., FROESE, G. & VIGILANT, L. (2018) Long-term group membership and dynamics in a wild western lowland gorilla population (*Gorilla gorilla gorilla*) inferred using non-invasive genetics. *American Journal of Primatology* **80**. John Wiley and Sons Inc.
- HAIMOFF, E.H. (1986) Convergence in the duetting of monogamous Old World primates. *Journal of Human Evolution* **15**, 51–59.
- HAMMERSCHMIDT, K. & FISCHER, J. (2019) Baboon vocal repertoires and the evolution of primate vocal diversity. *Journal of Human Evolution* **126**, 1–13. Academic Press.
- HARCOURT, A. & STEWART, K. (1996) Function and Meaning of Wild Gorilla ‘Close’ Calls II. Correlations with Ranks and Relatedness. *Behaviour*, 827–845.
- HARCOURT, A.H. & STEWART, K.J. (2007) *Gorilla society: conflict, compromise, and cooperation between the sexes*. In *Chicago: The University of Chicago Press*. p. Chicago, IL: The University of Chicago Press.
- HARDUS, M.E., LAMEIRA, A.R., SINGLETON, I., MORROGH-BERNARD, H.C., KNOTT, C.D., ANCRENAZ, M., UTAMI ATMOKO, S.S. & WICH, S.A. (2009) A description of the orangutan’s vocal and sound repertoire, with a focus on geographic variation. In *Orangutans: Geographic Variation in Behavioral Ecology and Conservation* (ed OXFORD UNIVERSITY PRESS), pp. 49–64. New-York.
- HARE, B., MELIS, A.P., WOODS, V., HASTINGS, S. & WRANGHAM, R. (2007) Tolerance Allows Bonobos to Outperform Chimpanzees on a Cooperative Task. *Current Biology* **17**, 619–623.
- HARE, B. & TOMASELLO, M. (2004) Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. *Animal Behaviour* **68**, 571–581.
- HEDWIG, D., MUNDRY, R., ROBBINS, M.M. & BOESCH, C. (2015a) Audience effects, but not environmental influences, explain variation in gorilla close distance vocalizations-A test of the acoustic adaptation hypothesis. *American Journal of Primatology* **77**, 1239–1252.

- HEDWIG, D., MUNDY, R., ROBBINS, M.M. & BOESCH, C. (2015b) Contextual correlates of syntactic variation in mountain and western gorilla close-distance vocalizations: Indications for lexical or phonological syntax? *Animal Cognition* **18**, 423–435.
- HENRY, L., CRAIG, A.J.F.K., LEMASSON, A. & HAUSBERGER, M. (2015) Social coordination in animal vocal interactions. Is there any evidence of turn-taking? The starling as an animal model. *Frontiers in Psychology* **6**, 1416.
- HIRATA, S. & FUWA, K. (2007) Chimpanzees (*Pan troglodytes*) learn to act with other individuals in a cooperative task. *Primates* **48**, 13–21.
- HOBOLTH, A., DUTHEIL, J.Y., HAWKS, J., SCHIERUP, M.H. & MAILUND, T. (2011) Incomplete lineage sorting patterns among human, chimpanzee, and orangutan suggest recent orangutan speciation and widespread selection. *Genome Research* **21**, 349–356.
- HOHMANN, G. & FRUTH, B. (1994) Structure and use of distance calls in wild bonobos (*Pan paniscus*). *International Journal of Primatology* **15**, 767–782.
- HOLLE, H. & GUNTER, T.C. (2007) The Role of Iconic Gestures in Speech Disambiguation: ERP Evidence. *Journal of Cognitive Neuroscience* **19**, 1175–1192.
- VAN HOOFF, J.A.R.A.M. & VAN SCHAİK, C.P. (1994) Male Bonds: Affiliative Relationships Among Non-human Primate Males. *Behaviour*, 309–337.
- IDANI, G., MWANZA, N., IHOBE, H., HASHIMOTO, C., TASHIRO, Y. & FURUICHI, T. (2008) Changes in the status of bonobos, their habitat, and the situation of humans at Wamba in the Luo Scientific Reserve, Democratic Republic of Congo. In *The bonobos: behavior, ecology, and conservation* (ed J. FURUICHI, T. & THOMPSON), pp. 121–134. Springer, New York.
- JANIK, V.M. (2000) Whistle Matching in Wild Bottlenose Dolphins (*Tursiops truncatus*). *Science* **289**, 1355–1357.
- JENSEN, K., HARE, B., CALL, J. & TOMASELLO, M. (2006) What's in it for me? Self-regard precludes altruism and spite in chimpanzees. *Proceedings of the Royal Society B: Biological Sciences* **273**, 1013–1021. Royal Society.
- KANO, T. (1982) The Social Group of Pygmy Chimpanzees (*Pan paniscus*) of Wamba. *PRIMATES* **23**, 171–188.
- KANO, T. (1992) *The last ape: Pygmy chimpanzee behavior and ecology*. Stanford University Press, Stanford, California.
- KAPPELER, P.M. & VAN SCHAİK, C.P. (2002) Evolution of Primate Social Systems. *International Journal of Primatology* **23**, 707–740.
- KEENAN, S., LEMASSON, A. & ZUBERBÜHLER, K. (2013) Graded or discrete? A quantitative analysis of Campbell's monkey alarm calls. *Animal Behaviour* **85**, 109–118.
- KEENAN, S., MATHEVON, N., STEVENS, J.M.G., NICOLÈ, F., ZUBERBÜHLER, K., GUÉRY, J.-P. & LEVRÉRO, F. (2020) The reliability of individual vocal signature varies across the bonobo's graded repertoire. *Animal Behaviour* **169**, 9–21.
- KENDON, A. (2004) *Gesture: Visible Action as Utterance*. Cambridge University Press, Cambridge.
- KERSHENBAUM, A., FREEBERG, T.M. & GAMMON, D.E. (2015) Estimating vocal repertoire size is like collecting coupons: A theoretical framework with heterogeneity in signal abundance. *Journal of Theoretical Biology* **373**, 1e11.
- KNOTT, C., BEAUDROT, L., SNAITH, T., WHITE, S., TSCHAUNER, H. & PLANANSKY, G. (2008) Female-female competition in Bornean orangutans. *International Journal of Primatology* **29**, 975–997.

- KODA, H. (2004) Flexibility and context-sensitivity during the vocal exchange of coo calls in wild Japanese macaques (*Macaca fuscata yakui*). *Behaviour* **141**, 1279–1296.
- KODA, H., LEMASSON, A., OYAKAWA, C., RIZALDI, PAMUNGKAS, J. & MASATAKA, N. (2013) Possible Role of Mother-Daughter Vocal Interactions on the Development of Species-Specific Song in Gibbons. *PLoS ONE* **8**.
- KUMMER, H. (1971) *Primate Societies: Group Techniques of Ecological Adaptation*. Aldine, Chicago.
- KURODA, S. (1980) Social Behavior of the Pygmy Chimpanzees. *PRIMATES* **21**, 181–197.
- LAMEIRA, A.R. & CALL, J. (2020) Understanding Language Evolution: Beyond Pan-Centrism. *BioEssays*, 1900102.
- LANGERGRABER, K.E., MITANI, J.C., WATTS, D.P. & VIGILANT, L. (2013) Male-female socio-spatial relationships and reproduction in wild chimpanzees. *Behavioral Ecology and Sociobiology* **67**, 861–873.
- VAN LAWICK-GOODALL, J. (1968) The Behaviour of Free-living Chimpanzees in the Gombe Stream Reserve. *Animal Behaviour Monographs* **1**, 165–311.
- LEHMANN, J. & BOESCH, C. (2008) Sexual differences in chimpanzee sociality. *International Journal of Primatology* **29**, 65–81.
- LEHMANN, J. & BOESCH, C. (2009) Sociality of the dispersing sex: the nature of social bonds in West African female chimpanzees, Pan troglodytes. *Animal Behaviour* **77**, 377–387.
- LEHMANN, J., KORSTJENS, A.H. & DUNBAR, R.I.M. (2007) Group size, grooming and social cohesion in primates. *Animal Behaviour* **74**, 1617–1629.
- LEMASSON, A. (2011) What can forest guenons « tell » us about the origin of language? In *Primate Communication and Human Language: Vocalisation, gestures, imitation and deixis in humans and non-humans* (eds A. VILAIN, A., SCHWARTZ, J. L. & J. C., VAUCLAIR), pp. 39–70. John Benjamins Publishing Company, Amsterdam.
- LEMASSON, A., GANDON, E. & HAUSBERGER, M. (2010a) Attention to elders' voice in non-human primates. *Biology Letters* **6**, 325–328.
- LEMASSON, A., GLAS, L., BARBU, S., LACROIX, A., GUILLOUX, M., REMEUF, K. & KODA, H. (2011a) Youngsters do not pay attention to conversational rules: Is this so for non-human primates? *Scientific Reports* **1**.
- LEMASSON, A. & HAUSBERGER, M. (2004) Patterns of vocal sharing and social dynamics in a captive group of Campbell's monkeys (*Cercopithecus campbelli campbelli*). *Journal of Comparative Psychology* **118**, 347–359.
- LEMASSON, A. & HAUSBERGER, M. (2011) Acoustic variability and social significance of calls in female Campbell's monkeys ( *Cercopithecus campbelli campbelli* ). *The Journal of the Acoustical Society of America* **129**, 3341–3352. Acoustical Society of America (ASA).
- LEMASSON, A., HAUSBERGER, M. & ZUBERBÜHLER, K. (2005) Socially meaningful vocal plasticity in adult Campbell's monkeys (*Cercopithecus campbelli*). *Journal of Comparative Psychology* **119**, 220–229.
- LEMASSON, A., OUATTARA, K., BOUCHET, H. & ZUBERBÜHLER, K. (2010b) Speed of call delivery is related to context and caller identity in Campbell's monkey mal. *Naturwissenschaften* **97**, 1–5.
- LEMASSON, A., OUATTARA, K., PETIT, E.J. & ZUBERBÜHLER, K. (2011b) Social learning of vocal structure in a non-human primate? *BMC Evolutionary Biology* **11**, 1–7.
- LEMASSON, A., PEREIRA, H. & LEVRÉRO, F. (2018) Social Basis of Vocal Interactions in Western

- Lowland Gorillas (*Gorilla g. gorilla*). *Journal of Comparative Psychology* **132**, 141–151.
- LEONG, K.M., ORTOLANI, A., BURKS, K.D., MELLEN, J.D. & SAVAGE, A. (2003) Quantifying Acoustic and Temporal Characteristics of Vocalizations for a Group of Captive African Elephants *Loxodonta Africana*. *Bioacoustics The International Journal of Animal Sound and its Recording* **13**, 213–231.
- LEVINSON, S.C. (2016) Turn-taking in Human Communication – Origins and Implications for Language Processing. *Trends in Cognitive Sciences* **20**, 6–14.
- LEVINSON, S.C. & HOLLER, J. (2014) The origin of human multi-modal communication. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**. Royal Society of London.
- LEVRERO, F. (2005) Structure d'une population de gorilles (*Gorilla g. gorilla*) visitant une clairière forestière: nature et rôle des rencontres intergroupes dans sa dynamique. . Rennes.
- LEVRÉRO, F., CLERC, O., TOUITOU, S., FRÉDET, J., GUÉRY, J.-P. & LEMASSON, A. (2017) Bonobos converse and pay attention to breaking conversational rules. In *Behaviour 2017 - 35th International Ethological Conference* (ed S. E DA V. (ISPA); A. FOR THE S. OF A.B. (ASAB) INSTITUTO UNIVERSITÁRIO DE CIÊNCIAS PSICOLÓGICAS), p. hal-01576988. Estoril, Portugal.
- LEVRÉRO, F., GATTI, S., MÉNARD, N., PETIT, E., CAILLAUD, D. & GAUTIER-HION, A. (2006) Living in nonbreeding groups: An alternative strategy for maturing gorillas. *American Journal of Primatology* **68**, 275–291. John Wiley and Sons Inc.
- LEVRÉRO, F., TOUITOU, S., FRÉDET, J., NAIRAUD, B., GUÉRY, J.-P. & LEMASSON, A. (2019) Social bonding drives vocal exchanges in Bonobos. *Scientific Reports* **9**, 711.
- LIEBAL, K., PIKA, S.S. & TOMASELLO, M. (2006) Gestural communication of orangutans (*Pongo pygmaeus* ). *Gesture* **6**, 1–38.
- MACKINNON, J. & MACKINNON, K. (1980) The behavior of wild spectral tarsiers. *International Journal of Primatology* **1**, 361–379.
- MAISELS, F., STRINDBERG, S., BREUER, T., GREER, D., JEFFERY, K. & STOKES, E. (2018) Gorilla gorilla ssp. gorilla (amended version of 2016 assessment). The IUCN Red List of Threatened Species 2018.
- MANSER, M.B., JANSEN, D.A.W.A.M., GRAW, B., HOLLÉN, L.I., BOUSQUET, C.A.H., FURRER, R.D. & LE ROUX, A. (2014) Vocal Complexity in Meerkats and Other Mongoose Species. In *Advances in the Study of Behavior* pp. 281–310. Academic Press Inc.
- MARLER, P. & TENAZA, R. (1977) Signaling behavior of apes with special reference to vocalizations. In *How Animals Communicate* (ed T.A. SEBEOK), pp. 965–1033. Indiana University Press, Bloomington.
- MASATAKA, N. & BIBEN, M. (1987) Temporal Rules Regulating Affiliative Vocal Exchanges of Squirrel Monkeys. *Behaviour* **101**, 311–319.
- MCCOMB, K. & SEMPLE, S. (2005) Coevolution of vocal communication and sociality in primates. *Biology Letters* **1**, 381–385.
- MCCOMB, K.E. (1991) Female choice for high roaring rates in red deer, *Cervus elaphus*. *Animal Behaviour* **41**, 79–88.
- MCNEILL, D. (2000) *Language and Gesture*. Cambridge University Press, Cambridge.
- MELIS, A.P., HARE, B. & TOMASELLO, M. (2006) Engineering cooperation in chimpanzees: tolerance constraints on cooperation. *Animal Behaviour* **72**, 275–286.



- MITANI, J.C. (2009a) Cooperation and competition in chimpanzees: Current understanding and future challenges. *Evolutionary Anthropology* **18**, 215–227.
- MITANI, J.C. (2009b) Male chimpanzees form enduring and equitable social bonds. *Animal Behaviour* **77**, 633–640.
- MITANI, J.C. & BRANDT, K.L. (1994) Social Factors Influence the Acoustic Variability in the Long Distance Calls of Male Chimpanzees. *Ethology* **96**, 233–252.
- MITANI, J.C. & NISHIDA, T. (1993) Contexts and social correlates of long-distance calling by male chimpanzees. *Animal Behaviour* **45**, 735–746.
- MITANI, J.C., WATTS, D.P. & AMSLER, S.J. (2010) Lethal intergroup aggression leads to territorial expansion in wild chimpanzees. *Current Biology* **20**, R507–R508.
- MORROGH-BERNARD, H.C., MORF, N. V., CHIVERS, D.J. & KRÜTZEN, M. (2011) Dispersal Patterns of Orang-utans (*Pongo* spp.) in a Bornean Peat-swamp Forest. *International Journal of Primatology* **32**, 362–376.
- MORTON, E.S. (1977) On the occurrence and significance of motivational-structural rules in some bird and mammal sounds. *The American Naturalist* **111**, 855–869.
- MULLER, M.N. & MITANI, J.C. (2005) Conflict and Cooperation in Wild Chimpanzees. *Advances in the Study of Behavior* **35**, 275–331.
- NAGUIB, M. & KIPPER, S. (2006) Effects of different levels of song overlapping on singing behaviour in male territorial nightingales (*Luscinia megarhynchos*). *Behavioral Ecology and Sociobiology* **59**, 419–426.
- NAGUIB, M. & MENNILL, D.J. (2010) The signal value of birdsong: empirical evidence suggests song overlapping is a signal. *Animal Behaviour* **80**, 11–15.
- NAKAMURA, M. (2000) Is Human Conversation more efficient than Chimpanzee Grooming? Comparison of Clique Sizes. *Human Nature* **11**, 281–297.
- NAKAMURA, M. (2003) 'Gathering' of social grooming among wild chimpanzees: Implications for evolution of sociality. *Journal of Human Evolution* **44**, 59–71. Academic Press.
- NARAT, V., PENNEC, F., SIMMEN, B., NGAWOLO, J.C.B. & KRIEF, S. (2015) Bonobo habituation in a forest–savanna mosaic habitat: influence of ape species, habitat type, and sociocultural context. *Primates* **56**, 339–349. Springer-Verlag Tokyo.
- NATER, A., NIETLISBACH, P., ARORA, N., VAN SCHAIK, C.P., VAN NOORDWIJK, M.A., WILLEMS, E.P., SINGLETON, I., WICH, S.A., GOOSSENS, B., WARREN, K.S., VERSCHOOR, E.J., PERWITASARI-FARAJALLAH, D., PAMUNGKAS, J. & KRÜTZEN, M. (2011) Sex-biased dispersal and volcanic activities shaped phylogeographic patterns of extant orangutans (genus: *Pongo*). *Molecular Biology and Evolution* **28**, 2275–2288.
- NISHIDA, T. (1968) The social group of wild chimpanzees in the Mahali Mountains. *Primates* **9**, 167–224.
- VAN NOORDWIJK, M.A., ARORA, N., WILLEMS, E.P., DUNKEL, L.P., AMDA, R.N., MARDIANAH, N., ACKERMANN, C., KRÜTZEN, M. & VAN SCHAIK, C.P. (2012) Female philopatry and its social benefits among Bornean orangutans. *Behavioral Ecology and Sociobiology* **66**, 823–834.
- OATES, J.F., TUTIN, C.E.G., HUMLE, T., WILSON, M.L., BAILLIE, J.E.M., BALMFORTH, Z., BLOM, A., BOESCH, C., COX, D., DAVENPORT, T., DUNN, A., DUPAIN, J., DUVALL, C., ELLIS, C.M., FARMER, K.H., ET AL. (2007) Pan troglodytes: IUCN Red List of Threatened Species.
- OWREN, M.J. & RENDALL, D. (1997) An Affect-Conditioning Model of Non-human Primate Vocal Signaling. In *Perspective in Ethology, Volume 12* (ed OWINGS ET AL.), pp. 299–346 Plenum

Press. New York.

- PAPWORTH, S., BÖSE, A.-S., BARKER, J., SCHEL, A.-M. & ZUBERBÜHLER, K. (2008) Male blue monkeys alarm call in response to danger experienced by others. *Biology Letters* **4**, 472–475.
- PARNELL, R.J. (2002) Group size and structure in western lowland gorillas (*Gorilla gorilla gorilla*) at Mbeli Bai, Republic of Congo. *American Journal of Primatology* **56**, 193–206.
- PATTERSON, N., RICHTER, D.J., GNERRE, S., LANDER, E.S. & REICH, D. (2006) Genetic evidence for complex speciation of humans and chimpanzees. *Nature* **441**, 1103–1108. Nature Publishing Group.
- PELÉ, M., DUFOUR, V., THIERRY, B. & CALL, J. (2009) Token Transfers Among Great Apes (*Gorilla gorilla*, *Pongo pygmaeus*, *Pan paniscus*, and *Pan troglodytes*): Species Differences, Gestural Requests, and Reciprocal Exchange. *Journal of Comparative Psychology* **123**, 375–384. American Psychological Association Inc.
- PIKA, S., WILKINSON, R., KENDRICK, K.H. & VERNES, S.C. (2018) Taking turns: bridging the gap between human and animal communication. *Proceedings of the Royal Society B* **285**, 1–9.
- PISANSKI, K., CARTEI, V., MCGETTIGAN, C., RAINE, J. & REBY, D. (2016) Voice Modulation: A Window into the Origins of Human Vocal Control? Elsevier Ltd. *Trends in Cognitive Sciences*. .
- POLLARD, K.A. & BLUMSTEIN, D.T. (2012) Evolving communicative complexity: Insights from rodents and beyond. Royal Society. *Philosophical Transactions of the Royal Society B: Biological Sciences*. .
- POOLE, T.B. (1987) Social behavior of a group of orangutans (*Pongo pygmaeus*) on an artificial island in Singapore Zoological Gardens. *Zoo Biology* **6**, 315–330.
- POUGNAULT, L., LEVRÉRO, F. & LEMASSON, A. (2020a) Conversation among primate species. In *The Origins of Language Revisited - Differentiation from Music and the Emergence of Neurodiversity and Autism* (ed N. MASATAKA), p. in Press. Springer, Singapor.
- POUGNAULT, L., LEVRÉRO, F., MULOT, B. & LEMASSON, A. (2020b) Breaking conversational rules matters to captive gorillas: A playback experiment. *Scientific Reports* **10**, 6947.
- PRÜFER, K., MUNCH, K., HELLMANN, I., AKAGI, K., MILLER, J.R., WALENZ, B., KOREN, S., SUTTON, G., KODIRA, C., WINER, R., KNIGHT, J.R., MULLIKIN, J.C., MEADER, S.J., PONTING, C.P., LUNTER, G., ET AL. (2012) The bonobo genome compared with the chimpanzee and human genomes. *Nature* **486**, 527–531.
- RAVIGNANI, A., BOWLING, D. & FITCH, W.T. (2014) Chorusing, synchrony and the evolutionary functions of rhythm. *Frontiers in Psychology* **5**, 1118.
- REBOUT, N., DE MARCO, A., LONE, J.-C., SANNA, A., COZZOLINO, R., MICHELETTA, J., STERCK, E.H.M., LANGERMANS, J.A.M., LEMASSON, A. & THIERRY, B. (2020) Tolerant and intolerant macaques show different levels of structural complexity in their vocal communication. *Proceedings of the Royal Society B: Biological Sciences* **287**, 20200439.
- RIJKSEN, H.D. (1978) *field study on Sumatran Orang-Utans (Pongo pygmaeus abelii, lesson 1827): Ecology, Behaviour, and Conservation*.
- ROBBINS, M.M., BERMEJO, M., CIPOLLETTA, C., MAGLIOCCA, F., PARNELL, R.J. & STOKES, E. (2004) Social structure and life-history patterns in western gorillas (*Gorilla gorilla gorilla*). *American Journal of Primatology* **64**, 145–159.
- ROBERTS, S.G.B. & ROBERTS, A.I. (2019) Social and ecological complexity is associated with gestural repertoire size of wild chimpanzees. *Integrative Zoology* **n/a**, n/a. Wiley.
- ROGERS, J. & GIBBS, R.A. (2014) Comparative primate genomics: Emerging patterns of genome

- content and dynamics. *Nature Reviews Genetics* **15**, 347–359. Nature Publishing Group.
- SAKAMAKI, T. (2013) Social grooming among wild bonobos (*Pan paniscus*) at Wamba in the Luo Scientific Reserve, DR Congo, with special reference to the formation of grooming gatherings. *Primates* **54**, 349–359.
- SALMI, R., HAMMERSCHMIDT, K. & DORAN-SHEEHY, D.M. (2013) Western gorilla vocal repertoire and contextual use of vocalizations. *Ethology* **119**, 831–847.
- VAN SCHAIK, C.P., MARSHALL, A.J. & WICH, S.A. (2009) Geographic variation in orangutan behavior and biology. Its functional interpretation and its mechanistic basis. In *Orangutans: Geographic variation in behavioral ecology and conservation* (ed C.P. WICH, S.A., UTAMI-ATMOKO, S.S., SETIA, T.M., AND VAN SCHAIK), pp. 351–361. Oxford University Press, New York.
- SHEL, A.M. & ZUBERBÜHLER, K. (2012) Dawn chorusing in guereza colobus monkeys. *Behavioral Ecology and Sociobiology* **66**, 361–373.
- SCHERER, K.R. (2003) Vocal communication of emotion: A review of research paradigms. *Speech Communication* **40**, 227–256.
- SCHULZ, T.M., WHITEHEAD, H., GERO, S. & RENDELL, L. (2008) Overlapping and matching of codas in vocal interactions between sperm whales: insights into communication function. *Animal Behaviour* **76**, 1977–1988.
- SEWALL, K.B. (2015) Social complexity as a driver of communication and cognition. In *Integrative and Comparative Biology* pp. 384–395. Oxford University Press.
- SHULTZ, S. & DUNBAR, R.I.M. (2007) The evolution of the social brain: Anthropoid primates contrast with other vertebrates. *Proceedings of the Royal Society B: Biological Sciences* **274**, 2429–2436. Royal Society.
- SINGLETON, I., KNOTT, C.D., MORROGH-BERNARD, H.C., WICH, S.A. & VAN SCHAIK, C.P. (2009) Ranging behavior of orangutan females and social organization. In *Orangutans: Geographic Variation in Behavioral Ecology and Conservation* p.
- SLOCOMBE, K. & ZUBERBÜHLER, K. (2010) Vocal Communication in Chimpanzees. In *The Mind of the Chimpanzee: Ecological and Experimental Perspectives* p. University of Chicago Press, Chicago & London.
- SNOWDON, C.T. & CLEVELAND, J. (1984) “Conversations” among pygmy marmosets. *American Journal of Primatology* **7**, 15–20.
- SNOWDON, C.T. & ELOWSON, A.M. (1999) Pygmy marmosets modify call structure when paired. *Ethology* **105**, 893–908.
- STERN, D.N., JAFFE, J., BEEBE, B. & BENNETT, S.L. (1975) Vocalizing in unison and in alternation: Two modes of communication within the mother- infant dyad. *Annals of the New York Academy of Sciences* **263**, 89–100.
- STIVERS, T., ENFIELD, N.J., BROWN, P., ENGLERT, C., HAYASHI, M., HEINEMANN, T., HOYMANN, G., ROSSANO, F., DE RUITER, J.P., YOON, K.-E. & LEVINSON, S.C. (2009) Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences* **106**, 10587–10592.
- STOKES, E.J. (2004) Within-group social relationships among females and adult males in wild western lowland gorillas (*Gorilla gorilla gorilla*). In *American Journal of Primatology* pp. 233–246.
- STOKES, E.J., PARNELL, R.J. & OLEJNICZAK, C. (2003) Female dispersal and reproductive success in wild western lowland gorillas (*Gorilla gorilla gorilla*). *Behavioral Ecology and Sociobiology* **54**, 329–339.

- SUCHAK, M., EPPLEY, T.M., CAMPBELL, M.W., FELDMAN, R.A., QUARLES, L.F. & DE WAAL, F.B.M. (2016) How chimpanzees cooperate in a competitive world. *Proceedings of the National Academy of Sciences* **113**, 10215–10220.
- SUGIURA, H. (1993) Temporal and Acoustic Correlates in Vocal Exchange of Coo Calls in Japanese Macaques. *Behaviour* **124**, 3–4.
- SUGIURA, H. & MASATAKA, N. (1995) Temporal and acoustic flexibility in vocal exchanges of coo calls in Japanese macaques (*Macaca fuscata*). In *Current topics in primate vocal communication* (eds E. ZIMMERMANN, J.D. NEWMAN & U. JÜRGENS), pp. 121–140. NY : Plenum Press. New York.
- SURBECK, M., GIRARD-BUTTOZ, C., BOESCH, C., CROCKFORD, C., FRUTH, B., HOHMANN, G., LANGERGRABER, K.E., ZUBERBÜHLER, K., WITTIG, R.M. & MUNDRY, R. (2017) Sex-specific association patterns in bonobos and chimpanzees reflect species differences in cooperation. *Royal Society Open Science* **4**, 1–20. Royal Society.
- TAKAHASHI, D.Y., FENLEY, A.R. & GHAZANFAR, A.A. (2016) Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**, 20150370.
- TAKAHASHI, D.Y., NARAYANAN, D.Z. & GHAZANFAR, A.A. (2013) Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys. *Current Biology* **23**, 2162–2168.
- UPHAM, N.S., ESSELSTYN, J.A. & JETZ, W. (2019) Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. *PLoS Biology* **17**. Public Library of Science.
- DE WAAL, F.B.M. (1988) The Communicative Repertoire of Captive Bonobos (*Pan Paniscus*), Compared To That of Chimpanzees. *Behaviour* **106**, 183–251.
- DE WAAL, F.B.M. (1995) Primates: sex and societies. Collective matriarchat in Bonobo. *Sci Am* **272**, 82–88.
- WASER, P.M. (1977) Individual Recognition, Intragroup Cohesion and Intergroup Spacing: Evidence From Sound Playback to Forest Monkeys. *Behaviour* **60**, 28–73.
- WATTS, D.P. & MITANI, J.C. (2001) Boundary Patrols and Intergroup Encounters in Wild Chimpanzees. *Behaviour* **138**, 299–327.
- WELCH, G.F. (2005) Singing as communication. In *Musical Communication* (ed & D.H. D. MIELL, R.MACDONALD), pp. 239–259. Oxford University Press, New York.
- WILLIAMSON, E.A. & BUTYNSKI, T.M. (2013) Gorilla gorilla Western Gorilla. In *Mammals of Africa* (ed J.K. AND J.K. T.M. BUTYNSKI), pp. 39–45. Bloomsbury Publishing, London, UK.
- WILSON, M.L., HAUSER, M.D. & WRANGHAM, R.W. (2001) Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Animal Behaviour* **61**, 1203–1216.
- WILSON, M.L., KAHLBERG, S.M., WELLS, M. & WRANGHAM, R.W. (2012) Ecological and social factors affect the occurrence and outcomes of intergroup encounters in chimpanzees. *Animal Behaviour* **83**, 277–291. Academic Press.
- WITTIG, R.M., CROCKFORD, C., SEYFARTH, R.M. & CHENEY, D.L. (2007) Vocal alliances in Chacma baboons (*Papio hamadryas ursinus*). *Behavioral Ecology and Sociobiology* **61**, 899–909.
- WRANGHAM, R.W. (1999) Evolution of Coalitionary Killing. *Yearbook of Physical Anthropology* **42**, 1–30.
- WRANGHAM, R.W. & SMUTS, B.B. (1980) Sex differences in the behavioural ecology of chimpanzees

in the Gombe National Park, Tanzania. *Journal of Reproduction and fertility Supplement*, 13–31.

YOSHIDA, S. & OKANOYA, K. (2005) Evolution of turntaking: a bio-cognitive perspective. *Cognitive Studies* **12**, 153–165.

## **“Conversational rules” are driven by social pressures, say studies on great apes.**

Loïc Pougnault<sup>1,2,7</sup>, Florence Levréro<sup>2\*</sup>, Maël Leroux<sup>3,4,5</sup>, Julien Paulet<sup>1</sup>, Pablo Bombani<sup>6</sup>, Fabrice Dentressangle<sup>2</sup>, Laure Deruti<sup>2</sup>, Baptiste Mulot<sup>7</sup>, Alban Lemasson<sup>1,8\*</sup>

<sup>1</sup> Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000 Rennes, France.

<sup>2</sup> Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES / CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France.

<sup>3</sup> Department of Comparative Linguistics, University of Zürich, Zürich, Switzerland.

<sup>4</sup> Budongo Conservation Field Station, Masindi, Uganda.

<sup>5</sup> Center for the Interdisciplinary Study of Language Evolution (ISLE), University of Zürich, Zürich, Switzerland.

<sup>6</sup> NGO Mbou-Mon-Tour (Republic Democratic of Congo)

<sup>7</sup> ZooParc de Beauval & Beauval Nature, 41110 Saint Aignan, France.

<sup>8</sup> Institut Universitaire de France, Paris, France.

\* Same contribution

## **Supplementary materials**

Given the lack of literature concerning the vocal interactions in Bornean orang-utans, we conducted a study on two captive populations of orang-utans using the same approach as already applied by our team on other non-human great apes (gorillas: Lemasson, Pereira, & Levréro, 2018; bonobos: Levréro *et al.*, 2019; chimpanzees: Pougnault *et al.* Submitted) in a way to highlight potential vocal interaction patterns in this species.

### **Compliance with ethical standards**

All applicable international, national, and/or institutional guidelines for the care and use of animals were followed. The study has been conducted in accordance with the current laws in France (agreement with the 2010/63/UE). Animal husbandry and care were under the management of the animal caregivers of the ZooParc de Beauval, France. Our study was based on non-invasive observations and was approved by the ethics committee “Comité Rennais

d'Ethique en matière d'Expérimentation Animale" (*i.e.* Rennes Ethical comity for experiments using animals; CNREEA 2018041710224608-APAFIS#14920).

### ***Group compositions***

The study was conducted on two captive populations of Bornean orang-utans housed in the ZooParc de Beauval (France, afterwards ZPB) and in the Ouwehand Direnpark (Netherlands, afterwards OD). Both populations were composed of eight individuals but their distributions in sub-groups differed. In ZPB, all individuals (one adult flanged male, three adult females, two adolescent males, and two infants; age classes based on Wich *et al.*, 2004, see Table 1) were housed together in an indoor enclosure (165m<sup>2</sup>, 5.30m high) with an access to a 1000m<sup>2</sup> outdoor enclosure. Orang-utans from OD were distributed in four sub-groups according to their kinship (1: one flanged adult male, 2: one adult female and one adolescent male, 3: one adult female and one infant, 4: one adult female and two infants, see Table 1). Each sub-group lived in an indoor enclosure (~110m<sup>2</sup>x7m) with an audio-visual access to all other conspecifics, and an alternating access to a 750m<sup>2</sup> outdoor enclosure throughout the day. All adult females were unrelated and all immature individuals were sired by the current adult male of each zoo. All indoor enclosures were composed of glass walls, wired roof, straw litter and wood perches and temperature and humidity controlled. All outdoor enclosures were composed of grass, bush and wood perches. Food – composed by vegetables, fruits and dry food – was distributed 6 times a day and water was available *ad libitum*.

**Table 1** | Composition of the studied groups

Subject name (birth date – arrival date)	Age-sex category	Mother's name (all fathered by the adult male of the group)	Zoo	Sub-group number
Muda (29/05/1983 – 20/07/1994)	Adult flanged male		ZPB	
Christina (1985 <sup>a</sup> – 26/11/1993)	Adult female		ZPB	
Manis (24/09/1990 – 13/04/2010)	Adult female		ZPB	
Janah (15/02/1991 – 18/06/1992)	Adult female		ZPB	
Dayak (04/03/2009)	Adolescent male	Manis	ZPB	
Belayan (22/09/2009)	Adolescent male	Janah	ZPB	
Sabah (06/08/2015)	Infant female	Manis	ZPB	
Kayan (06/05/2017)	Infant male	Janah	ZPB	
Bako (21/03/1996)	Adult flanged male		OD	1
Anak (09/03/1976)	Adult female		OD	2
Tjintah (01/05/1984 – 19/08/2005)	Adult female		OD	3
Jewel (1987 <sup>a</sup> – 13/04/1992)	Adult female		OD	4
Jingga (22/10/2007)	Adolescent male	Jewel	OD	2
Daya (27/05/2012)	Infant female	Jewel	OD	4
Sabar (11/05/2016)	Infant male	Jewel	OD	4
Minggu (10/03/2019)	Infant female	Tjintah	OD	3

<sup>a</sup> Uncertainty regarding the date of birth of these individuals, estimation  $\pm 1$  year.

### ***Vocal data collection***

Call production of the two groups were collected by one observer (L.P.) on an *all occurrence* basis (Altmann, 1974) using a Marantz PMD660 handheld digital audio recorder (Marantz, Japan) (sample rate 44.1kHz, resolution 32bits, wav. format) connected to a directional Sennheiser MKH70-1 microphone (Sennheiser, Germany), and to a directional AV-JEFE TCM 160 LAVALIER tie microphone (AV-JEFE, Singapore) for the experimenter's comments. The four sub-groups housed at OD could be recorded all together given the layout of the enclosures. Call recording were conducted during recording sessions lasting 1:30 hr randomly distributed throughout the day between 9 am and 6 pm, with a maximal distance of 10 meters. Orang-utans had been habituated to the presence of the observer and recording equipment for seven days before the data collection.

A total of 72 calls were recorded in 17 hrs between July to November 2018 (*i.e.* excluding nine weeks during summer to avoid the season with the highest visitor presence rate) at ZPB, and 398 calls were collected (only one call was discarded due to the uncertainty in caller identity) in 25 hrs of recording in November 2019 at OD (only one call was discarded due to the



uncertainty in caller identity). To determine the call types, we referred to the previously published vocal repertoire done by Hardus *et al.* (2009). Table 2 provides the distribution of the different call types recorded in this study.

**Table 2** | Distribution of call types recorded in this study

		Ahh	Bark	Cry	Fast Long Call	Grunt	Grumph	Grumble	Kiss squeak
ZPB	Adult flanged male	0	0	0	3	1	0	1	0
	Adult females	0	0	0	0	0	0	0	0
	Adolescent males	0	0	0	0	1	0	0	0
	Infants	0	0	55	0	0	11	0	0
OD	Adult flanged male	0	0	0	4	29	0	2	0
	Adult females	0	2	0	0	60	0	92	3
	Adolescent male	7	0	0	0	116	0	10	0
	Infants	0	0	61	0	1	10	0	0

### ***Vocal data processing***

We investigated the temporal organisation of the call production within each population by measuring all the inter-calls durations based on spectrograms generated by PRAAT software version 6.1.09 using a Gaussian window. The inter-calls duration corresponded to the silence gap between two consecutive calls of any call types emitted by two different identified callers, from the offset of the first call to the onset of the second call (as in Lemasson *et al.*, 2018). We defined a temporal window of a maximum of 10 seconds after the onset of the first call, which was large enough regarding other findings in multiple species to highlight an eventual pattern of vocal interaction (Campbell's monkeys Lemasson, Gandon, *et al.*, 2010; Gorillas Lemasson *et al.*, 2018 and Bonobos Levréro *et al.*, 2019).

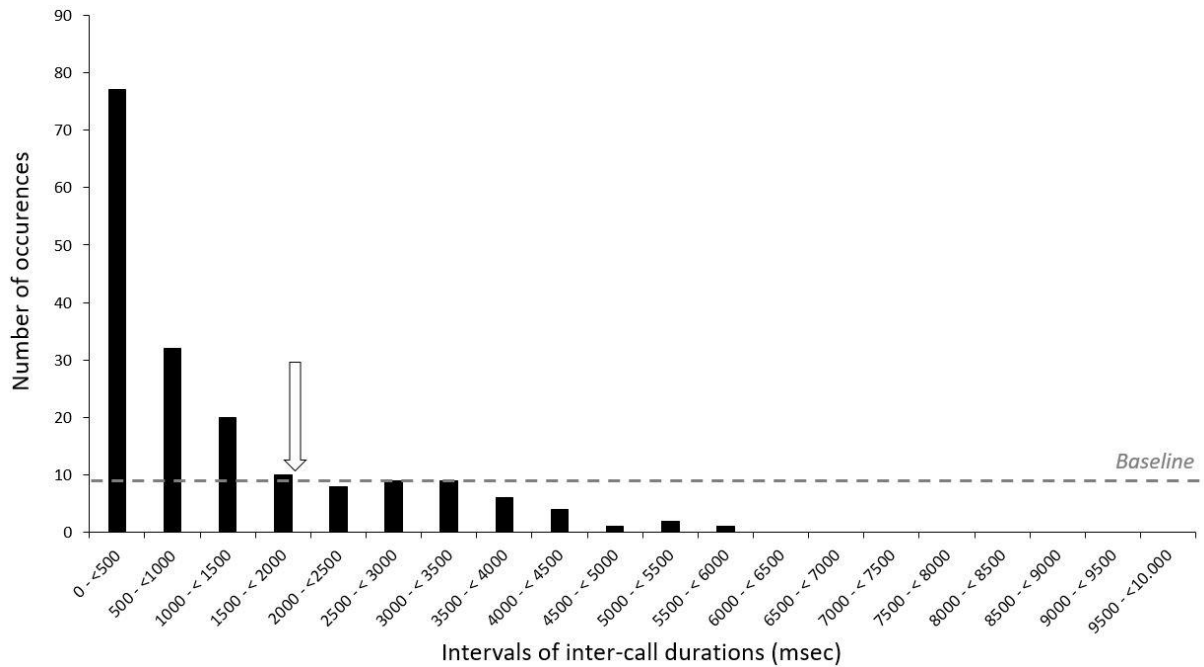
We were thus able to quantify the proportions of: (i) call utterances produced during vocal interactions, and (ii) call utterances emitted in absence of vocal interactions, which consisted of calls emitted independently from the vocal utterances of other group members (see also Bouchet, Pellier, Blois-Heulin, & Lemasson, 2010; Levréro *et al.*, 2019).

Also, call utterances from one individual emitted out of vocal interactions can be emitted following two temporal calling patterns: a call alone or a same call type repeated and separated by short silent gaps. For clarity, these two temporal calling patterns were respectively named: (i) “Isolated Call”, (ii) “Repeated Call”. To define if a call had to be considered as Repeated or Isolated, we used the same method as for vocal interactions (*i.e.* temporal window of a maximum of 10 seconds after the onset of the first call), but applied to inter-call durations of involving single callers. We then represented the distribution of inter-call delays and defined the temporal threshold to classify whether a call belonged to a series of Repeated Calls or was emitted as an Isolated Call. The threshold indicated the maximum delay for call repetition. It was placed at the location where inter-call delays start rising above a baseline representing a background noise of inter-call delay out of call repetition. We investigated the proportion of Isolated and Repeated Calls in both studied populations.

Lastly, we compared the hourly production rates of the different calling patterns according to the age-sex classes in both populations.

## ***Results***

From all the recorded calls (N= 469), only 2 calls from two different individuals occurred in a time window of 10 s with an inter-call delay of 1.3 s, the remaining 447 calls uttered out of the determined 10s-time window and the second shorter inter-call delay was 53 s. At the individual level, 179 inter-call delays (n= 358 calls) were recorded with a duration shorter than a 10s-time window. This allowed to graphically defined an inter-call delay of a maximum of 2 s between two successive calls from the same emitter to define if a call was Repeated or not (Figure 1). Then, Isolated Call represented 11.1% (8/72) of the total call production at ZPB and 49.6% (197/397) at OD, while Repeated Calls represented 88.9% (64/72) of the total call production at ZPB and 50.4% (200/397) at OD.



**Figure1** | *Distribution of inter-call durations between two successive calls from single callers recorded from all occurrence samples (n = 179 inter-call durations recorded in a time window of 10 s after hearing a call) in two zoos (ZooParc de Beauval – France – and Ouwehand Direnpark – Netherlands –). Grey dotted line corresponds to the baseline of the occurrence of positive inter-call durations. The white arrow corresponds to the threshold indicating the maximum response delay for Repeated calling, and was placed at the location where inter-call delays start rising above the baseline when moving from long delays to the peak of short delays.*

The hourly rate of call production was higher in the group from the OD than in the ZPB, respectively  $1.9 \pm 1.9$  and  $0.7 \pm 1.9$ . At ZPB, infants were the most vocally active individuals and produced all the Repeated Calls, while the hourly rate of call production was higher for the adolescent male at OD concerning the Isolated and Repeated Calls (see Table 3).

**Table 3** | Rate of vocal production per hour for each age-sex category of callers from ZPB and OD populations.

		Adult flanged males	Adult females	Adolescent males	Infants
All Calls	ZooParc de Beauval	0.29	0	0.029±0.042	1.9±0.92
	Ouwehand Direnpark	1.4	2.1±2.4	5.3	0.96±0.46
Isolated Calls	ZooParc de Beauval	0.29	0	0.029±0.042	0.041±0.33
	Ouwehand Direnpark	1.1	1.3±1.2	4.0	0.2±0.11
Repeated Calls	ZooParc de Beauval	0	0	0	1.5±0.67
	Ouwehand Direnpark	0.28	0.84±1.4	1.28	0.76±0.36

## Conclusion

Our results show a clear absence of vocal interactions, regardless of the living social conditions, in captive Bornean orang-utans. Nonetheless, population of OD was more vocally active than ZPB one. This difference may be due to the visual access of the individuals to the caregivers' kitchen. At OD, orang-utans highly solicit caregivers using gestures or grunt/grumble vocalisations in a way to obtain food or to start a medical training implying food rewards.

## References

- ALTMANN, J. (1974) Observational Study of Behavior: Sampling Methods. *Behaviour* **49**, 227–267.
- BOUCHET, H., PELLIER, A.S., BLOIS-HEULIN, C. & LEMASSON, A. (2010) Sex differences in the vocal repertoire of adult red-capped mangabeys (*Cercocebus torquatus*): A multi-level acoustic analysis. *American Journal of Primatology* **72**, 360–375.
- HARDUS, M.E., LAMEIRA, A.R., SINGLETON, I., MORROGH-BERNARD, H.C., KNOTT, C.D., ANCRENAZ, M., UTAMI ATMOKO, S.S. & WICH, S.A. (2009) A description of the orangutan's vocal and sound repertoire, with a focus on geographic variation. In *Orangutans: Geographic Variation in Behavioral Ecology and Conservation* (ed OXFORD UNIVERSITY PRESS), pp. 49–64. New-York.
- KODA, H. (2004) Flexibility and context-sensitivity during the vocal exchange of coo calls in wild Japanese macaques (*Macaca fuscata yakui*). *Behaviour* **141**, 1279–1296.
- LEMASSON, A., GANDON, E. & HAUSBERGER, M. (2010) Attention to elders' voice in non-human primates. *Biology Letters* **6**, 325–328.
- LEMASSON, A., GUILLOUX, M., RIZALDI, BARBU, S., LACROIX, A. & KODA, H. (2013) Age- and sex-dependent contact call usage in Japanese macaques. *Primates* **54**, 283–291.
- LEMASSON, A., PEREIRA, H. & LEVRÉRO, F. (2018) Social Basis of Vocal Interactions in Western Lowland Gorillas (*Gorilla g. gorilla*). *Journal of Comparative Psychology* **132**, 141–151.
- LEVRÉRO, F., TOUITOU, S., FRÉDET, J., NAIRAUD, B., GUÉRY, J.-P. & LEMASSON, A. (2019) Social bonding drives vocal exchanges in Bonobos. *Scientific Reports* **9**, 711.
- WICH, S.A., UTAMI-ATMOKO, S.S., MITRA SETIA, T., RIJKSEN, H.D., SCHÜRMAN, C., VAN HOOFF, J.A.R.A.M. & VAN SCHAIK, C.P. (2004) Life history of wild Sumatran orangutans (*Pongo abelii*). *Journal of Human Evolution* **47**, 385–398. Academic Press.

## **“Conversational rules” are driven by social pressures, say studies on great apes.**

Loïc Pougault<sup>1,2,7</sup>, Florence Levréro<sup>2\*</sup>, Maël Leroux<sup>3,4,5</sup>, Julien Paulet<sup>1</sup>, Pablo Bombani<sup>6</sup>, Fabrice Dentressangle<sup>2</sup>, Laure Deruti<sup>2</sup>, Baptiste Mulot<sup>7</sup>, Alban Lemasson<sup>1,8\*</sup>

<sup>1</sup> Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000 Rennes, France.

<sup>2</sup> Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES / CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France.

<sup>3</sup> Department of Comparative Linguistics, University of Zürich, Zürich, Switzerland.

<sup>4</sup> Budongo Conservation Field Station, Masindi, Uganda.

<sup>5</sup> Center for the Interdisciplinary Study of Language Evolution (ISLE), University of Zürich, Zürich, Switzerland.

<sup>6</sup> NGO Mbou-Mon-Tour (Republic Democratic of Congo)

<sup>7</sup> ZooParc de Beauval & Beauval Nature, 41110 Saint Aignan, France.

<sup>8</sup> Institut Universitaire de France, Paris, France.

\* Same contribution

## **Supplementary materials**

In a way to confirm the scarcity of vocal turn-taking in chimpanzees that we have already observed in a captive group (Pougault et al., submitted), and that Arcadi (2000) previously evidenced with wild adult males, we conducted a study on a free-ranging population using the same approach as already applied by our team on other non-human great apes (Gorillas: Lemasson, Pereira, & Levréro, 2018; Bonobos: Levréro *et al.*, 2019).

## **Compliance with ethical standards**

Permission and ethical approval to conduct the study was received from the Uganda Wildlife Authority (UWA) (permit number: COD/96/05), the Uganda National Council for Science and Technology (UNCST) (permit number: NS47ES) and the Animal Welfare & Ethical Review Body from the University of Warwick, United Kingdom (permit number: not provided)

## Group composition

The study was conducted between February to May 2019 on free-ranging chimpanzees of the Sonso community in the Budongo Forest Reserve (Uganda) containing 42 adult individuals (12 males and 30 females) on a home range of approximately 15km<sup>2</sup>. The Sonso community has been studied regularly since 1990 and is well habituated to human presence.

Given the difficulties to identify individuals in their natural habitat, our study included 35 well-identified individuals (12 adult males, 13 adult females, three sub adult males, one sub adult female, one juvenile male, one juvenile female, two infant males and two infant females; age-sex classes based on Reynolds, 2005, see Table 1).

**Table 1** | Composition of the identified individual of the studied community

Subject name (estimated birth date)	Age-sex category	Mother's name
NB (1962)	Adult female	-
OK (1996±1 year)	Adult female	-
FL (1979)	Adult female	-
KL (1979±1 year)	Adult female	-
KU (1979)	Adult female	-
KW (1981)	Adult female	-
ZF (1982±6 months)	Adult male	BN
JN (1984±1 year)	Adult female	-
TJ (1984±5 year)	Adult female	-
SQ (1991±1 year)	Adult male	BC
MS (1992±6 months)	Adult male	NB
HW (1993±1 year)	Adult male	HT
KT (1993)	Adult male	KU
SM (1993±1 year)	Adult male	SE
KZ (1995)	Adult male	KW
ZL (1995)	Adult male	ZN
PS (1 year)	Adult male	PL
RS (1997)	Adult female	RH
FK (1999±1 year)	Adult male	FL
RM (2002)	Adult female	RH
ZD (2003)	Adult male	ZN
DL (2002±1 year)	Adult female	-
CD (2003±1 year)	Adult female	-
KS (2003)	Adult male	KU
EV (2004±1 year)	Adult female	-
JS (2006)	Sub-adult male	JN
KC (2006)	Sub-adult male	KL
KX (2007)	Sub-adult female	KY
MB (2009)	Sub-adult male	ML
TW (2010)	Juvenile female	TJ
KF (2014)	Juvenile male	KU
KO (2014)	Infant male	KL
IS (2017)	Infant female	IR
RN (2018)	Infant male	RS
DB (2018)	Infant female	DL

- Unknown individual.

### **Vocal data collection**

Call productions were collected by one observer (ML), accompanied by a field assistant, using a Marantz PMD661 mk3 handheld digital audio recorder (Marantz, Japan) (sample rate 44.1kHz, resolution 32bits, .wav format) connected to a directional Sennheiser ME66/K6 microphone (Sennheiser, Germany). Call recordings were conducted on a *focal sampling* basis for adult individuals and in *ad libitum* for immature ones (*i.e.* sub adults, juveniles and infants) (Altmann, 1974) throughout the day between 7 am and 4:30 pm. Recording sessions lasted 2hrs per day, with a maximal distance of 15 meters (10 meters on average, sometimes at a greater distance when individuals were moving through the canopy).

The 1305 calls recorded corresponded to 44:34 hours of recording. Due to a too poor quality (e.g. too far emitters, or overlapping forest noises), 3% (41/1305) of calls were discarded. Among the remaining calls, 44% (562/1264) were also removed from our analyses given the uncertainty in caller identity. So, we conserved only 702 calls.

Given that turn-taking and avoidance of overlap are expected to be associated with calm and affiliation, while high arousal and conflicts lead to frequent overlap (e.g. Levréro *et al.*, 2019; Pournault, Levréro, & Lemasson, 2020), we distinguished calls emitted within peaceful contexts (e.g. resting, foraging periods, and affiliative social interactions such as grooming) from calls recorded in agonistic interactions (e.g. conflicts characterized by chasing, hitting and biting behaviours) or given in response to external disturbances (e.g. falling tree). Thus, from the initial 702 calls, we extracted 71% (495/702) of calls produced in peaceful contexts. To determine call types, we used the same criterion as in Pournault *et al.* (Submitted) (see Table 2).

**Table 2** | Distribution of call types recorded in this study.

	Adult females	Adult males	Sub-adult female	Sub-adult males	Juvenile females	Juvenile males	Infants
Bark	0	10	0	0	0	0	0
Cough	0	1	0	0	0	0	0
Grunt	220	112	0	11	0	4	3
Huu	0	0	0	0	0	0	0
Laugh	0	0	0	0	0	0	0
Mixed-calls	5	44	0	4	0	0	0
Pant	4	0	0	1	0	0	0
Pant-Grunt	22	15	1	6	2	1	1
Pant-Hoot	11	26	1	1	0	1	0
Rough Grunt	0	3	0	0	0	0	0
Scream	18	18	0	7	0	0	16
Squeak	17	0	0	0	0	0	0
Waa-Bark	2	0	0	0	0	0	0
Whimper	5	49	0	0	0	0	30
Hoo	14	13	0	3	0	0	0

### ***Vocal data processing***

All following analyses were performed both for calls emitted in all contexts ( $N_{\text{All contexts}} = 702$ ) and for calls only produced in peaceful contexts ( $N_{\text{Peaceful contexts}} = 495$ ).

First, we investigated the temporal organisation of the call production within the group measuring all the inter-call durations thanks to the spectrograms generated by PRAAT software version 6.1.09 using a Gaussian window. The inter-calls duration corresponded to the silence gap between two consecutive calls of any call types emitted by two different callers, from the offset of the first call to the onset of the second call (as in Lemasson et al., 2018). To define if a call belonged to a vocal interaction (*i.e.* which corresponded to at least two successive calls of alternating individuals with a short or a negative inter-calls delay) or not, we used a graphical method already applied to define the maximum response delay for a vocal interaction in several primate species (Campbell's monkeys: Lemasson, Gandon, et al., 2010; gorillas: Lemasson et



al., 2018 and bonobos: Levréro et al., 2019). We defined a temporal window of a maximum of 10 seconds after the onset of the first call, which was large enough to highlight an eventual pattern of vocal interaction if we refer to comparative studies on a broad range of species. We then represented the distribution of inter-call delays and defined the temporal threshold to classify whether a call belonged to a vocal interaction or not. The threshold indicated the maximum response delay for vocal interaction. It was placed at the location where inter-call delays start rising above a baseline representing a background noise of inter-calls delay out of vocal interactions. Given our analytic method, choruses (*i.e.* more than two individuals calling at the same time, Yoshida & Okanoya, 2005) have to be deleted due to the hard identification of callers implied. We conserved only overlapped calls implying two identified callers.

When several calls overlapped (negative inter-call durations), we conserved only the calls where callers were identifiable.

We were thus able to quantify the proportions of: (i) call utterances produced during vocal interactions, and (ii) call utterances emitted in absence of vocal interaction, which consisted of calls emitted independently from the vocal utterances of other group members (see also Bouchet, Pellier, Blois-Heulin, & Lemasson, 2010; Levréro et al., 2019).

Call utterances from one individual emitted out of vocal interactions can be emitted following two temporal calling patterns: a call alone or a same call type repeated (*i.e.* series of calls separated by short silent gaps). The same temporal threshold than the one defined for vocal interactions was applied to classify whether a call was emitted alone or repeated (Koda, 2004; Lemasson et al., 2013, 2018). For clarity, these two temporal calling patterns were respectively named: (i) “Isolated Call”, (ii) “Repeated Call”. It is worth noting that in the unique cases of *pant*, *pant-hoot* and *pant-grunt* vocalizations, Isolated Calls could be composed of several unmerged sound units because *pants*, *pant-hoots* and *pant-grunts* are, by definition, multi-unit

calls incorporating noisy inhalation phases (Gros-Louis & Mitani, 1998; Laporte & Zuberbühler, 2010).

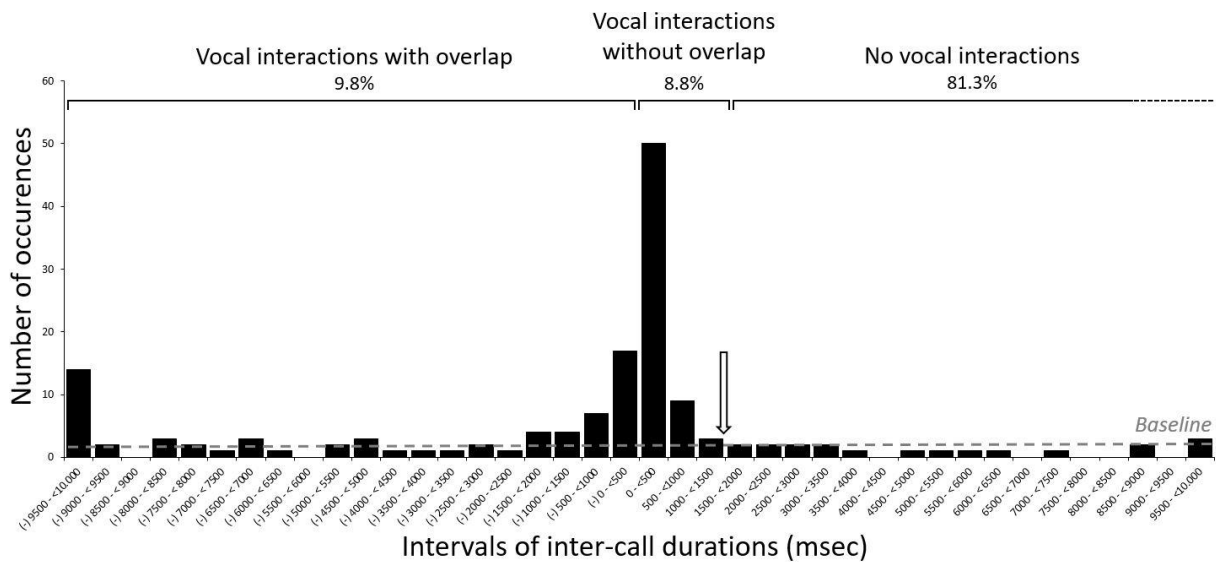
Lastly, we analysed the hourly production rates of the different calling patterns according to the sex of adult individuals.

## ***Results***

From all the recorded calls (N= 702), 150 calls from 35 individuals occurred in a time window of 10 seconds. From these 150 calls, we studied the distribution of inter-calls duration (n= 75 inter-calls duration). We graphically determined a positive inter-calls delay of a maximum of 1.5 s between two successive calls emitted by two different callers (Fig. 1) to define a vocal interaction. We thus found that only 18.7% (131/702) of calls belonged to vocal interactions.

Among the vocal interactions, we noticed some negative inter-calls durations that corresponded to call overlaps. Overlapped calls occurred in 52.7% (69/131) of the calls produced in vocal interactions. Calls produced out of vocal interactions (n=571) were composed by 11% (63/571) of Isolated Calls and 89% (508/571) of Repeated Calls.

If we considered only calls emitted in peaceful contexts (N=495), results are similar. We found that 7.9 % (39/495) of the total call production were produced in vocal interactions. Overlapped calls occurred in 64.1% (25/39), while vocal interactions without overlap represented 35.9% (14/39). Calls produced out of vocal interactions (n=456) were composed by 13.6% (62/456) of Isolated calls and 86.4% (394/456) of Repeated Calls.



**Figure 1** | Distribution of inter-calls durations between two successive calls from different callers recorded from all occurrence samples ( $n = 75$  inter-calls durations recorded in a time window of 10 seconds after hearing a call). Grey dotted line corresponds to the baseline of the occurrence of positive inter-calls durations. The white arrow corresponds to the threshold indicating the maximum response delay for vocal interaction, and was placed at the location where inter-calls delays start rising above the baseline when moving from long delays to the peak observed at around 1.5 s.

Adult males were in general more vocally active than adult females during vocal interactions, with an equivalent proportion of interactions with or without overlap in each sex. Isolated and Repeated Calls were equivalently produced by females and males in all contexts, with a higher production of Repeated Calls. In peaceful contexts, females were more vocally active than males out of vocal interactions, particularly concerning Repeated Calls.

**Table 3** | Rate of vocal production per hour of adult females and males both in all contexts and in peaceful contexts.

	All contexts					Peaceful contexts				
	All calls	Calls in vocal interactions		Calls out of vocal Interactions		All calls	Calls in vocal interactions		Calls out of vocal interactions	
		With overlap	Without overlap	Isolated	Repeated		With overlap	Without overlap	Isolated	Repeated
Adult females	2.3±3.6	0.08±0.11	0.10±0.14	0.24±0.31	1.9±3.3	2.1±3.6	0.05±0.06	0.08±1.0	0.23±0.32	1.7±3.4
Adult males	2.3±2.7	0.39±0.47	0.4±0.59	0.21±0.23	1.9±1.7	1.1±1.5	0.11±0.13	0.13±0.13	0.18±0.22	0.73±1.2

## Conclusion

Whether we considered all contexts or only peaceful contexts, four calling patterns were identified: call overlap, vocal response without overlap and a vast majority of calls emitted out of vocal interactions composed of Isolated and Repeated Calls. Overlapping calls represented the core of vocal interactions and vocal turn-taking was rare. Adult males were the most vocally active during vocal interactions in all contexts and only in peaceful context, while adult females were particularly vocal in terms of Isolated and Repeated Call in peaceful contexts. That confirms results recently obtained in captivity (Pougnault et al. Submitted).

## References

- ALTMANN, J. (1974) Observational Study of Behavior: Sampling Methods. *Behaviour* **49**, 227–267.
- BOUCHET, H., PELLIER, A.S., BLOIS-HEULIN, C. & LEMASSON, A. (2010) Sex differences in the vocal repertoire of adult red-capped mangabeys (*Cercocebus torquatus*): A multi-level acoustic analysis. *American Journal of Primatology* **72**, 360–375.
- GROS-LOUIS, J. & MITANI, J. (1998) Chorusing and call convergence in chimpanzees: Tests of three hypotheses. *Behaviour* **135**, 1041–1064.
- KODA, H. (2004) Flexibility and context-sensitivity during the vocal exchange of coo calls in wild Japanese macaques (*Macaca fuscata yakui*). *Behaviour* **141**, 1279–1296.
- LAPORTE, M.N.C. & ZUBERBÜHLER, K. (2010) Vocal greeting behaviour in wild chimpanzee females. *Animal Behaviour* **80**, 467–473. Elsevier Ltd.
- LEMASSON, A., GANDON, E. & HAUSBERGER, M. (2010) Attention to elders' voice in non-human primates. *Biology Letters* **6**, 325–328.

- LEMASSON, A., GUILLOUX, M., RIZALDI, BARBU, S., LACROIX, A. & KODA, H. (2013) Age- and sex-dependent contact call usage in Japanese macaques. *Primates* **54**, 283–291.
- LEMASSON, A., PEREIRA, H. & LEVRÉRO, F. (2018) Social Basis of Vocal Interactions in Western Lowland Gorillas (*Gorilla g. gorilla*). *Journal of Comparative Psychology* **132**, 141–151.
- LEVRÉRO, F., TOUITOU, S., FRÉDET, J., NAIRAUD, B., GUÉRY, J.-P. & LEMASSON, A. (2019) Social bonding drives vocal exchanges in Bonobos. *Scientific Reports* **9**, 711.
- POUGNAULT, L., LEVRÉRO, F. & LEMASSON, A. (2020) Conversation among primate species. In *The Origins of Language Revisited - Differentiation from Music and the Emergence of Neurodiversity and Autism* (ed N. MASATAKA), p. in Press. Springer.
- REYNOLDS, V. (2005) *The chimpanzees of the Budongo Forest*. Oxford University Press, Oxford.
- YOSHIDA, S. & OKANOYA, K. (2005) Evolution of turntaking: a bio-cognitive perspective. *Cognitive Studies* **12**, 153–165.

## “Conversational rules” are driven by social pressures, say studies on great apes.

Loïc Pougault<sup>1,2,7</sup>, Florence Levréro<sup>2\*</sup>, Maël Leroux<sup>3,4,5</sup>, Julien Paulet<sup>1</sup>, Pablo Bombani<sup>6</sup>, Fabrice Dentressangle<sup>2</sup>, Laure Deruti<sup>2</sup>, Baptiste Mulot<sup>7</sup>, Alban Lemasson<sup>1,8\*</sup>

<sup>1</sup> Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000 Rennes, France.

<sup>2</sup> Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES / CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France.

<sup>3</sup> Department of Comparative Linguistics, University of Zürich, Zürich, Switzerland.

<sup>4</sup> Budongo Conservation Field Station, Masindi, Uganda.

<sup>5</sup> Center for the Interdisciplinary Study of Language Evolution (ISLE), University of Zürich, Zürich, Switzerland.

<sup>6</sup> NGO Mbou-Mon-Tour (Republic Democratic of Congo)

<sup>7</sup> ZooParc de Beauval & Beauval Nature, 41110 Saint Aignan, France.

<sup>8</sup> Institut Universitaire de France, Paris, France.

\* Same contribution

## Supplementary materials

Knowledge of the production rate of vocal utterances of wild bonobos is lacking. Most of the available studies focused on the description of the bonobo's vocal repertoire (e.g in captivity: De Waal, 1988 ; *in situ*: Bermejo & Omedes, 1999) and on the acoustic structure and functions of specific call types (in captivity : e.g. Clay & Zuberbühler, 2009, 2011; Clay *et al.*, 2011; Clay, Archbold, & Zuberbühler, 2015; White *et al.*, 2015; Keenan *et al.*, 2020). Levréro and colleagues (2019) investigated the bonobos' vocal exchanges in the framework of conversation-like communication, but still in captivity. Recently there is a growing interest in intergroup vocal communication among wild sub-groups (Schamberg *et al.*, 2016, 2017), as a function of inter-party movement coordination. Indeed, bonobos live in fission-fusion society and form temporary sub-groups for hours or days whose composition changes through time (van Elsacker, Vervaecke, & Verheyen, 1995) and re-merge in different group compositions. Here, we propose to complete the general knowledge about vocal behaviour of wild bonobos by assessing their hourly vocal production rate. So far this information is totally unknown.

Through this supplementary note, we only focused on intra-sub-groups vocal behaviour and we reported the vocal production of each focused sub-group (one at a time).

### **Compliance with ethical standards**

All methods were carried out in accordance with the relevant guidelines and regulations and were approved by the Institutional Animal Ethical Committee of the University of Lyon/Saint-Etienne, under the authorization no. 42-218-0901-38 SV 09 (Lab ENES) and the host institutions WWF-DRC, WWF-Belgium and the NGO Mbou-Mon-Tour. This study only involves behavioural observations and spontaneous vocal recordings of animals in their natural social group and in their natural environment. No experimental protocols and invasive methods have been conducted in this study.

### ***Group composition***

Our data came from the free-ranging bonobo community of Nkala forest (2°36'10"S, 16°27'17"E), located in the Forest Concession of Local Communities of the Mbali River (previously called "Malebo"), in the Lac Tumba landscape, Mai-Ndombe province, Democratic Republic of Congo (DRC), co-led by the Congolese NGO Mbou-Mon-Tour (MMT) and WWF-DRC. The habitat of this area is composed of a forest-savannah mosaic with 60% of lowland tropical rainforest and 40 % of savannah (survey block of 1993 km<sup>2</sup>; Inogwabini et al. 2008). The studied group was composed of 19 identified individuals (six adult males, six adult females, and seven immatures: two young males, two young females and three individuals with unknown sex). The entire group may be a bit bigger since we observed up to 22 different individuals at the same time.

The study was conducted from January 27 to April 13 2012 (L.D. and F.L.) and from October 27 2012 to December 30 2012 (FD). The size of this group's home range was unknown. When we knew the location of the nest site from the night before, early in the next morning we waited at proximity of this site before they wake up. The group usually splitted into several sub-groups and we followed the larger sub-group. The average size of monitored subgroups was 11,3 individuals, ranging from three to 22.

### ***Vocal data collection***

All the call productions of the monitored sub-group were collected *ad libitum* in forest from 5:30 a.m. to 6 p.m. (average time in forest 4hrs49 / day) by one researcher (F.D., L.D. or F.L. depending on the period), accompanied by one to three field assistants, using a field notebook and a SONY voice recorder according to we were ourselves moving to follow the bonobos. Audio recordings of calls were not systematic done since the bonobos were often in movement making our own progression in forest and good-quality recordings very difficult. We noted all the call productions of bonobos that we could hear at a distance ranged of 15m to nearly 100m. In total, we were in visual and/or audio contact with the bonobos 64 days (91hrs24min) out of 98 days (473hrs) spent in forest. Over the 64 days, we heard the bonobos calling 56 days. We spent on average 1hrs18min by day in contact with the bonobos at all possible times of the observation day.

### ***Vocal data processing***

We first isolated so-called vocal events from our observations, independently of the number of callers. A vocal event corresponded to an isolated call, namely no other call was produced within the next three seconds, or a series of calls produced within three seconds of each other. When the number of calls emitted in series exceed five, we stopped counting them (see Fig1).



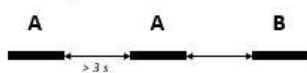
Two distinct vocal events were thus separated by more than three seconds. This temporal threshold was intuitively chosen in the field but actually matches with the temporal threshold that has been found in captive bonobos from the analysis of the distribution of all inter-calls delays (Levréro *et al.*, 2019).

Second, when events were composed of more than one call, we investigated for possible vocal interactions. Several individuals could indeed be involved in a single vocal event (Fig1). We estimated the number of different callers when we saw them calling or when they were distinctly located in space regarding the position of observers (the number of callers may be underestimated if two or more callers were physically at proximity).

Lastly, we focused on the calls from our focal sub-group only and did not include the calls from other sub-groups that could be heard (either in response of the calls of our focused group or spontaneous calls).

**3 vocal events:**

An unique call each time



**3 vocal events:**

1 vocal event with repeated calls (2 to 5 calls) produced by 1 emitter

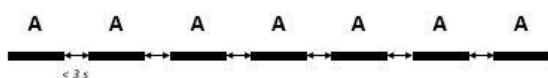
Followed by 1 vocal event produced by 3 emitters (2 to 5 calls)

Followed by 1 vocal event with repeated calls (2 to 5 calls) produced by 1 emitter



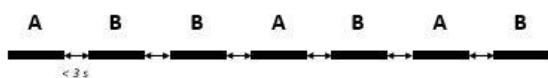
**1 vocal event:**

1 emitter, more than 5 calls



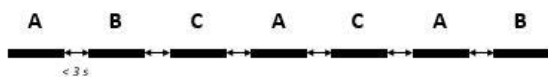
**1 vocal event:**

2 emitters, more than 5 calls



**1 vocal event:**

More than 2 emitters, more than 5 calls



Individuals : A, B, C Time line : from left to right

**Figure 1** | Representation of the different vocal events considered

**Results**

In total, 2166 vocal events were recorded and we discarded 34 events, which were due to external disturbances like our presence between two sub-groups willing to merge or gunshots in the neighbouring forest in order to not inflate artificially their call frequency rate. We thus conserved a total of 2132 vocal events.

Given the 91hrs24 spend in contact with bonobos, the hourly rate of vocal event calculated reached 23.31 vocal events per hour. Most of vocal events were produced by a single caller (74 %, see Tab.1). Vocal events produced by only two interlocutors were rare (5%), while 21% involved more than two callers. Most of vocal events were composed of several consecutive

calls (from 2 to 5 calls, 48%, see Table1), 29% were constituted by more than five calls and 16 % of a unique call.

**Table 1** | Distribution of vocal events according to the number of involved callers or the number of calls

Number of emitters				TOTAL vocal events
1	2	>2	No data available	
1579	104	346	103	2132
Number of calls				TOTAL vocal events
1	2 – 5	>5	No data available	
338	1013	624	157	2132

### Conclusion

Here we report the first assessment of the hourly vocal rate production in a free-ranging bonobo community. The tolerance of bonobos to our presence was very good (29 cases out 2132 vocal events were due to our presence) despite irregular presence of observers in forest through the year, which rules out an over estimation of calling rate due to a possible insufficient level of the habituation of this community.

On the contrary, because the vocal data collection was *ad libitum* since we could never be sure that we could detect all their occurrences, we may have under estimated the soft vocal interactions in peaceful context due to forest conditions and our minimum distance to the bonobos. Anyhow, bonobos appear highly vocally active, vocal events involving most of the time a single caller producing repeated calls.

### References

- BERMEJO, M. & OMEDES, A. (1999) Preliminary Vocal Repertoire and Vocal Communication of Wild Bonobos (*Pan paniscus*) at Lilungu (Democratic Republic of Congo). *Folia Primatologica* **70**, 328–357.
- CLAY, Z., ARCHBOLD, J. & ZUBERBÜHLER, K. (2015) Functional flexibility in wild bonobo vocal

- behaviour. *PeerJ* **3**, e1124.
- CLAY, Z., PIKA, S., GRUBER, T. & ZUBERBÜHLER, K. (2011) Female bonobos use copulation calls as social signals. *Biology Letters* **7**, 513–516.
- CLAY, Z. & ZUBERBÜHLER, K. (2009) Food-associated calling sequences in bonobos. *Animal Behaviour* **77**, 1387–1396.
- CLAY, Z. & ZUBERBÜHLER, K. (2011) Bonobos extract meaning from call sequences. *PLoS ONE* **6**, e18786.
- VAN ELSACKER, L., VERVAECKE, H. & VERHEYEN, R.F. (1995) A Review of Terminology on Aggregation Patterns in Bonobos (*Pan paniscus*). In *International Journal of Primatology* p.
- KEENAN, S., MATHEVON, N., STEVENS, J.M.G., NICOLÈ, F., ZUBERBÜHLER, K., GUÉRY, J.-P. & LEVRÉRO, F. (2020) The reliability of individual vocal signature varies across the bonobo's graded repertoire. *Animal Behaviour* **169**, 9–21.
- LEVRÉRO, F., TOUITOU, S., FRÉDET, J., NAIRAUD, B., GUÉRY, J.-P. & LEMASSON, A. (2019) Social bonding drives vocal exchanges in Bonobos. *Scientific Reports* **9**, 711.
- SCHAMBERG, I., CHENEY, D.L., CLAY, Z., HOHMANN, G. & SEYFARTH, R.M. (2016) Call combinations, vocal exchanges and interparty movement in wild bonobos. *Animal Behaviour* **122**, 109–116. Elsevier Ltd.
- SCHAMBERG, I., CHENEY, D.L., CLAY, Z., HOHMANN, G. & SEYFARTH, R.M. (2017) Bonobos use call combinations to facilitate inter-party travel recruitment. *Behavioral Ecology and Sociobiology* **71**, 75.
- DE WAAL, F.B.M. (1988) The Communicative Repertoire of Captive Bonobos (*Pan Paniscus*), Compared To That of Chimpanzees. *Behaviour* **106**, 183–251.
- WHITE, F.J., WALLER, M., BOOSE, K., MERRILL, M.Y. & WOOD, K.D. (2015) Function of loud calls in wild bonobos. *Journal of Anthropological Sciences* **93**, 1–13.



# Chapitre VI

## Discussion générale



© Peinture sur toile de Monique Brizard



## 1. Principaux résultats

La revue de littérature que nous avons réalisée (**article 1**) met en lumière les points communs entre les échanges vocaux organisés des primates non-humains et les conversations humaines. Lors de ces interactions, les individus respectent un « contrat de communication » reposant sur des règles temporelles et sociales. La première règle temporelle consiste à ne pas recouvrir la vocalisation de son interlocuteur par le respect d'un délai de réponse minimum. Ce court intervalle de silence permet de limiter les risques de coupure de parole, et d'assurer la fluidité de l'échange. La durée de ce délai peut également être modulé en accord avec des déterminants sociaux (*e.g.* délai plus court en réponse à la vocalisation d'un individu affilié). Le respect du tour de parole entre les interlocuteurs, qui se traduit par une alternance des vocalisations, constitue la seconde règle temporelle. Au cours de ces interactions, les interlocuteurs ne sont pas choisis au hasard ou de façon opportuniste, mais en fonction de caractéristiques sociales qui leurs sont propres (*i.e.* facteurs intrinsèques : âge, sexe ou relation de parenté ; facteurs extrinsèques : affinité sociale, rang hiérarchique). Enfin, des phénomènes de convergences vocales peuvent apparaître lors des échanges vocaux.

A l'amorce de ce projet de doctorat, les études se focalisaient sur les primates non-Hominidés et peu d'investigations avaient été réalisées chez les grands singes. Ce manque de données empiriques créait des lacunes majeures concernant la compréhension de l'importance sociale des échanges vocaux organisés, et la trajectoire évolutive de ce comportement vocal chez les primates. Les données les plus complètes à disposition provenaient d'études conduites chez les bonobos. En effet, des observations menées sur un groupe captif faisaient état d'échanges vocaux temporellement organisés (*i.e.* respect d'un délai inter-vocalisations court) fréquemment produits par des individus affiliés (Levréro *et al.* 2019). L'importance du respect de ces règles « conversationnelles », notamment l'évitement du recouvrement des signaux, a pu être confirmée expérimentalement chez cette espèce (Levréro *et al.* 2017). Une seconde série d'observations, chez les gorilles des plaines de l'Ouest cette fois, témoignait d'échanges vocaux fréquents entre individus d'âges proches respectant un court délai inter-vocalisations (Lemasson *et al.* 2018). Néanmoins, la confirmation expérimentale de la pertinence de ces possibles règles temporelles et sociales était essentielle (**article 2**). A l'inverse des résultats obtenus chez les bonobos et les gorilles, Arcadi (2000) soulignait l'absence d'échanges vocaux organisés chez les mâles chimpanzés de la communauté sauvage de Budongo (Ouganda). Ces résultats se devaient d'être complétés par des observations supplémentaires tenant compte de l'ensemble des individus (*i.e.* tout âge et sexe confondus) aussi bien en milieu captif (**article 3**)



qu'en milieu sauvage (**article 4**). Finalement, aucune information n'était disponible en ce qui concerne de possibles interactions vocales chez les orangs-outans. Il était donc crucial d'explorer cet aspect de leur communication vocale (**article 4**).

A la vue de l'état des connaissances, l'objectif de nos recherches était de compléter les informations parcellaires disponibles concernant les échanges vocaux organisés et les potentielles règles « conversationnelles » chez les grands singes non-humains (*i.e.* bonobo, chimpanzé, gorille des plaines de l'Ouest et orang-outan de Bornéo), d'évaluer la pertinence de ces règles pour les individus ainsi que l'influence de la vie sociale (*i.e.* structure et organisation sociales des groupes) sur les convergences et divergences répertoriées entre ces quatre espèces de primates. De manière plus générale, l'objectif de ce projet était de contribuer au débat théorique sur les possibles scénarios évolutifs de l'émergence de cette capacité langagière au sein de l'ordre de primates. Nos résultats ont permis de compléter les connaissances concernant les échanges vocaux organisés chez les primates Hominidés, tout en apportant des éléments de réflexion supplémentaire quant au développement de ce comportement vocal au cours de l'ontogénie, les subtilités des facteurs sociaux pouvant favoriser son émergence et son occurrence au sein de la lignée des primates.

Tout d'abord, nous démontrons par une approche expérimentale (**article 2**) la pertinence du respect du délai inter-vocalisations pour les gorilles des plaines de l'Ouest. Toutefois, nous ne confirmons que partiellement l'existence possible d'une règle sociale car la sensibilité au respect de la différence d'âge entre interlocuteurs dépend de l'âge même du sujet testé. De manière générale, l'attention portée aux échanges vocaux respectant ou non ces règles varie en fonction de l'âge du sujet ce qui suggère un possible rôle de l'apprentissage social.

Puis, nous confirmons l'extrême rareté des échanges vocaux chez les chimpanzés (**article 3** et **article 4**). Les interactions vocales sont dominées par les vocalisations fortes des mâles, produites en superposition (**article 3** et **article 4**). De plus, la majorité des émissions vocales des chimpanzés ne sont pas des interactions mais plutôt des émissions de cris répétés successivement par un même émetteur. Ce pattern d'émission est par ailleurs caractéristique d'un contexte de communication agonistique avec une influence du rang hiérarchique de l'individu sur son taux d'émission (**article 3**). Ces observations corroborent l'importance de la hiérarchie de dominance dans l'organisation sociale des chimpanzés (Goodall 1986; Müller et Mitani 2005) et questionnent l'absence d'échanges vocaux organisés chez cette espèce au regard de leur socialité particulièrement compétitive.

Enfin, l'approche comparative que nous avons conduite sur les quatre espèces de grands singes souligne que ces derniers utilisent des échanges vocaux basés sur des règles « conversationnelles », à l'image de l'humain et de beaucoup d'espèces de primates non-Hominidés (**article 4**). Nos résultats révèlent également l'importance de la vie sociale sur la structuration des interactions vocales car seules les espèces à la vie sociale riche interagissent vocalement. De plus, l'organisation temporelle de ces interactions vocales semble être associée à la qualité des liens sociaux existant entre les membres d'un groupe. L'émergence de la capacité à interagir vocalement sous la forme d'échanges vocaux organisés paraît médiée par des facteurs sociaux multiples traduisant une certaine complexité sociale tels que le niveau de tolérance et de compétition entre les membres du groupe. Ces observations suggèrent qu'au-delà de critères tels que la taille du groupe social ou le taux de toilettage social, il ne faut pas négliger la valence des relations sociales comme expression de la complexité sociale, pouvant sous-tendre à la complexité vocale.

Par la suite, nous discuterons des implications de nos résultats au regard des quatre questions proposées par Tinbergen (1963). Les données expérimentales obtenues chez les gorilles des plaines de l'Ouest (**article 2**) nous permettront d'aborder le développement ontogénétique de ce comportement vocal avant de traiter des fonctions supposées de celui-ci. Puis, en nous appuyant sur les résultats des observations menées chez les trois autres espèces (*i.e.* chimpanzé, bonobo et orang-outan de Bornéo : **article 3** et **article 4**), nous évoquerons les facteurs sociaux pouvant sous-tendre à l'émergence des échanges vocaux organisés pour finir par aborder la question de l'évolution de cette capacité langagière chez les primates. Les perspectives de recherche seront développées au terme de cette discussion, notamment concernant les précisions à apporter à de futurs protocoles et les questions de recherches pouvant être soulevées par nos recherches.

## 2. Développement ontogénétique des échanges vocaux

Le développement de la communication vocale chez les primates non-humains a longtemps été considéré comme étant sous l'emprise d'un déterminisme génétique fort (*e.g.* Hammerschmidt et Fischer 2008; Newman et Symmes 1982; Seyfarth et Cheney 1997). Sans négliger ces précédentes découvertes, d'autres travaux suggèrent que le développement vocal est également le résultat croisé de processus de maturation, d'apprentissage et d'expérience sociale menant à l'amélioration des capacités de compréhension, d'utilisation et, dans une

moindre mesure, de production des vocalisations (*e.g.* Boughman et Moss 2003; Janik et Slater 1997; Lemasson *et al.* 2011; Seyfarth et Cheney 1997).

Les résultats que nous avons obtenus auprès des jeunes gorilles des plaines de l'Ouest (**article 2**), couplés aux études conduites chez des espèces de primates non-Hominidés, semblent suggérer que la maîtrise (*i.e.* perception et utilisation) des règles « conversationnelles » n'est pas innée (*e.g.* mone de Campbel : Lemasson *et al.* 2010, 2011 ; macaque japonais : Bouchet *et al.* 2017; Lemasson *et al.* 2013a ; singe araignée *Ateles geoffroyi* : Briseño-Jaramillo *et al.* 2018 ; marmoset *Callithrix jacchus* : Chow *et al.* 2015; Takahashi *et al.* 2016).

A ce stade, deux hypothèses non exclusives peuvent être soulevées concernant les processus développementaux qui sous-tendent la maîtrise de ces règles temporelles et sociales qui régulent les échanges vocaux : (i) la maturation des substrats neuronaux sous-jacents aux mécanismes de perception et de contrôle moteur de la production des vocalisations et (ii) l'existence de mécanismes complexes basés sur l'apprentissage social de ces règles.

### 2.1. Contrôle moteur et développement des structures cérébrales

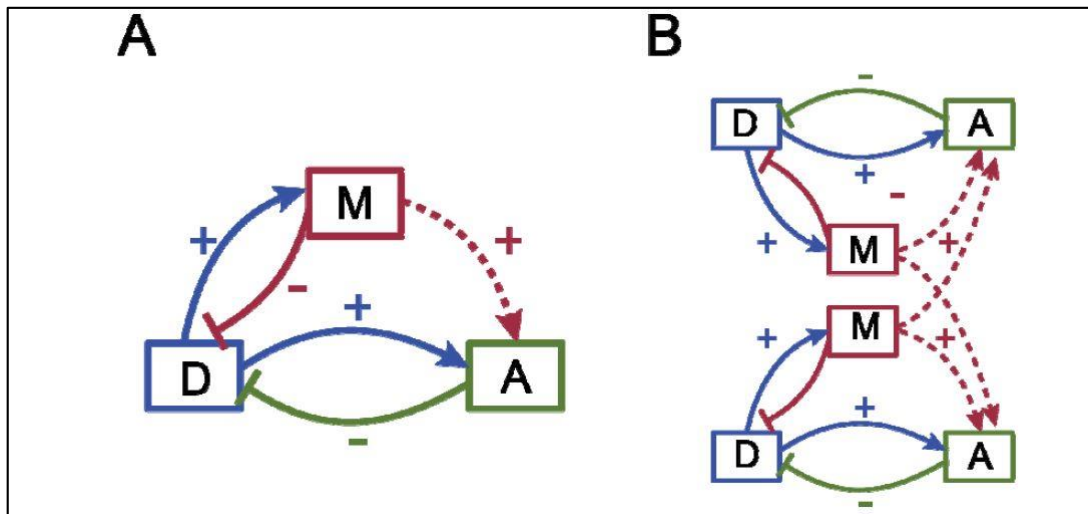
Des preuves s'accumulent quant au fait que les jeunes individus respectent moins que leurs aînés le délai inter-vocalisations lors des interactions vocales conversationnelles (*e.g.* Chow *et al.* 2015; Lemasson *et al.* 2013a), allant jusqu'à couper près de 12 fois plus la parole à leurs interlocuteurs que les adultes (Lemasson *et al.* 2010). Il en va de même chez les humains : bien que le tour de parole apparaissent très précocement dans notre espèce (Dominguez *et al.* 2016), dans les premières années de développement, les recouvrements de vocalisations sont encore fréquents (Ginsburg et Kilbourne 1988; Hilbrink *et al.* 2015).

En comparaison à l'espèce humaine (*e.g.* Fitch 2000; Ghazanfar et Rendall 2008), le contrôle vocal des primates non-humains a longtemps été considéré comme inexistant car lié au niveau d'excitation de l'individu et associé à des structures cérébrales sub-corticales, et notamment au système limbique (*e.g.* Aitken 1981; Jürgens 1995; Ploog 1981). Pourtant, les preuves d'un contrôle vocal non-négligeable s'accumulent grâce aux récentes études de conditionnement opérant. Il a ainsi été démontré qu'à la diffusion d'un bruit blanc, les primates sont en mesure d'interrompre leurs vocalisations (tamarin *Saguinus oedipus* : Miller *et al.* 2003; marmoset : Roy *et al.* 2011). Les primates sont également capables d'apprendre à vocaliser à la présentation d'un item alimentaire (macaque à queue de cochon *Macaca nemestrina* : Coudé *et al.* 2011), d'un stimulus visuel (macaque rhesus : Hage et Nieder 2013), ou encore sur commande en réponse à un geste de l'expérimentateur humain (gibbon à mains

blanches *Hylobates lar* : Koda *et al.* 2007). Aussi, ces investigations ont souligné le rôle des aires cérébrales corticales (*i.e.* cortex pré-moteur et cortex pré-frontal : Brumm *et al.* 2003; Coudé *et al.* 2011; Hage et Nieder 2013), également impliquées dans la production du langage chez l'humain (*i.e.* structures homologues à l'aire de Broca). Les interactions vocales sont généralement corrélées avec l'activation de ces mêmes aires corticales (*e.g.* marmoset : Nummela *et al.* 2017; Simões *et al.* 2010), suggérant une production volontaire des vocalisations dans ce contexte (*e.g.* Nummela *et al.* 2017; Simões *et al.* 2010; Toarmino *et al.* 2017). Nous pouvons alors supposer que le contrôle de la production vocale des primates non-humains nécessite un certain stade de maturation des systèmes cérébraux sous-jacents. Les individus les plus jeunes, dont l'anatomie cérébrale a été particulièrement peu étudiée jusqu'à présent, peuvent ne pas avoir atteint un stade de maturation cérébrale suffisant pour avoir un contrôle total de leurs productions vocales lors des interactions. Aussi, il est possible de supposer que les jeunes se distinguent des adultes par leur état affectif au moment de l'interaction. Ce manque d'auto-régulation peut alors entraîner chez ceux-ci un taux important de vocalisations répétées et de coupures de parole (Lemasson *et al.* 2013a).

Toutefois, il se peut que la difficulté des jeunes à respecter le tour de parole soit aussi due à une difficulté de perception des vocalisations de leurs congénères. En ce sens, une piste de réflexion peut être fournie par le modèle proposé par Takahashi et collaborateurs (2012, 2013a). Selon ces auteurs, le mécanisme cérébral sous-jacent au développement du contrôle des productions lors des échanges vocaux est constitué de trois structures neuronales - les régions limbique, motrice et auditive - impliquant des phénomènes de rétrocontrôles activateurs et inhibiteurs (Takahashi *et al.* 2012). Ce modèle s'articule en quatre étapes majeures (Figure 1A) : (i) l'activité du système limbique s'accroît lentement jusqu'à atteindre un seuil, (ii) il active alors le système moteur qui induit la production de la vocalisation, (iii) le signal acoustique produit est perçu puis active le système neuronal auditif (iv) qui va alors exercer une réduction progressive de l'activité du système limbique et permettre d'inhiber la production vocale (Takahashi *et al.* 2012, 2013b). À côté de cette dynamique centrale, les auteurs supposent l'existence d'une seconde boucle de rétroaction rapide dans laquelle l'activité du système limbique augmente par pics et active directement les régions motrices et auditives (Figure 1A). Cette seconde dynamique permettrait de générer des syllabes dans la vocalisation (Takahashi *et al.* 2012, 2013b). Lors d'un échange vocal impliquant deux partenaires (Figure 1B), le système auditif de l'interlocuteur est activé par le signal acoustique produit par

l'émetteur. Par cette activation, la production vocale de l'interlocuteur est inhibée et empêche ainsi la superposition des vocalisations des deux partenaires.



**FIGURE 1** | (A) Représentation schématique du modèle isolé montrant les interactions positives (+) et négatives (-) entre les différents substrats neuronaux (D = système limbique, M = système moteur, et A = système auditif). Les lignes pleines représentent le transfert d'informations par le biais des connexions neuronales et les lignes pointillées représentent le transfert d'informations par le milieu extérieur (*i.e.* l'air). (B) Représentation du fonctionnement du modèle lors d'un échange vocal entre deux individus. Source : issue de Takahashi *et al.* 2012.

En simulant ce modèle et en réduisant l'influence inhibitrice de la région auditive, les auteurs observent une augmentation du nombre de superpositions vocales (Takahashi *et al.* 2012). Ils en déduisent donc que le manque de contrôle des jeunes individus et le taux important de recouvrement de vocalisations qu'ils induisent découlent probablement d'un système neuronal auditif immature (Takahashi *et al.* 2012, 2013ab, 2016).

Des résultats similaires chez l'humain corroborent l'idée selon laquelle le manque de contrôle de la production vocale chez les plus jeunes serait probablement due à l'immaturité des systèmes neuronaux sous-jacents (MacDonald *et al.* 2012). Pourtant, il est actuellement difficile de savoir si ce manque de contrôle des productions vocales lors de l'échange résulte principalement d'une immaturité du système moteur ou du système auditif. Les méthodes d'imagerie fonctionnelle (TEP-IRM), qui ont déjà mis en évidence l'activité de certaines zones cérébrales lors de la production de gestes communicatifs chez les chimpanzés (Tagliabata *et al.* 2008), peuvent être des pistes intéressantes pour décrire les mécanismes neuronaux à

l'œuvre lors de la production d'échanges vocaux temporellement organisés et leur développement au cours de l'ontogénie.

## 2.2. Développement sous influences sociales

L'importance de l'influence sociale sur le développement de la communication vocale fait l'objet d'investigations depuis plusieurs dizaines d'années, et ce, particulièrement chez les espèces aviaires. Il a été démontré qu'en l'absence d'adultes, les juvéniles produisent des vocalisations « aberrantes » ne ressemblant pas à celles d'individus élevés dans un environnement social naturaliste (e.g. étourneau : Bertin *et al.* 2007; Apodiforme : Jarvis 2004; carouge à épaulettes *Agelaius phoeniceus* : Marler *et al.* 1972; diamant mandarin *Taeniopygia guttata* : Price 1979). Par ces expériences de privation sociale, les auteurs démontrent à quel point la présence d'un adulte comme modèle est crucial au bon développement vocal du jeune. Chez les primates non-humains, l'influence de l'environnement social sur le développement vocal est sujet à controverse (Bouchet *et al.* 2016). Des expériences d'isolement social, ou d'assourdissement à la naissance, suggèrent que les vocalisations des primates non-humains se manifestent dès la naissance avec une structure comparable à celle des adultes (Newman et Symmes 1974; Winter *et al.* 1973). Par ailleurs, à la suite d'hybridations et d'adoption croisées, les individus issues de ces unions émettaient des vocalisations à la structure intermédiaire entre les deux espèces parents (gibbon à mains blanches et gibbon à bonnet *Hylobates pileatus* : Brockelman et Schilling 1984; Geissmann 1984; macaque rhesus et macaque japonais : Owren *et al.* 1992), laissant supposer un fort déterminisme génétique. Toutefois, il s'avère que certaines vocalisations du répertoire vocal n'apparaissent pas sous leur structure adulte à la naissance, mais subissent un certain degré de changement au cours du développement ontogénétique (Elowson *et al.* 1992; Hauser 1989). Chez les marmosets (marmoset pygmé *Cebuella pygmaea* et marmoset commun), des phénomènes de « babillages » comparables à ce qui peut être observé chez l'humain (Marler et Tenaza 1977), ont été rapportés (Pistorio *et al.* 2006; Snowdon et Elowson 2001; Takahashi *et al.* 2015). Ainsi, au cours du développement, les longues séquences vocales des jeunes se raccourcissent et la structure de leurs vocalisations évolue vers des formes plus matures, notamment grâce à l'influence des adultes (Takahashi *et al.* 2015).

Cette influence sociale cruciale pour le développement de certains comportements vocaux chez les primates non-humains peut laisser penser que des mécanismes similaires sont à l'œuvre en ce qui concerne la maîtrise des échanges vocaux organisés au cours du développement ontogénétique chez ces derniers.

### 2.2.1. Apprentissage social des règles d'interactions

Pour participer de façon optimale aux échanges vocaux organisés temporellement et socialement, le jeune doit apprendre à produire ses vocalisations avec la temporalité adéquate pour respecter le tour de parole et éviter la superposition vocale, mais également à utiliser les vocalisations appropriées au contexte de l'interaction. Ces capacités d'usage, de perception et de compréhension émergent progressivement au cours du développement (*e.g.* Janik et Slater 1997).

Bien que probablement sous-tendue par des processus de maturation neuronale, la maîtrise de la temporalité des interactions vocales peut être favorisée par les interactions répétées avec les congénères, et tout particulièrement les parents. Par exemple, Chow et collaborateurs (2015) rapportent que chez les marmosets, la vitesse de réponse d'un jeune lors des interactions vocales « conversationnelles » avec l'un de ses parents a tendance à décroître au cours des huit ou neuf premiers mois de développement de ce dernier, réduisant de ce fait la probabilité de voir apparaître des superpositions vocales entre les interlocuteurs. Les auteurs suggèrent que cette réduction progressive des superpositions résulterait d'un renforcement induit par le parent : l'adulte ignorerait les vocalisations superposées du jeune et répondrait préférentiellement aux vocalisations produites avec la bonne temporalité (Chow *et al.* 2015). Si les échanges vocaux entre un jeune et l'un des parents ralentissent progressivement, la temporalité des interactions reste toutefois relativement constante entre les juvéniles d'une même portée (Chow *et al.* 2015). Ces résultats suggèrent que les changements ontogénétiques ne sont pas simplement dus à un processus de maturation globale (Takahashi *et al.* 2016), mais révèlent une forme d'apprentissage de la temporalité de production des vocalisations sensible au contexte social de production (Chow *et al.* 2015). Chez d'autres espèces de primates, tels que les gibbons, il semble que les interactions mère-enfant jouent un rôle important dans le développement vocal des jeunes, notamment en ce qui concerne le contrôle temporel des productions vocales rythmiques. En effet, au cours des six premières années de vie du jeune, une amélioration de la synchronisation temporelle au cours des duos se manifeste (Koda *et al.* 2013). De même chez l'humain, la mère et l'enfant semblent tous deux activement impliqués dans le processus d'acquisition de la régulation temporelle des vocalisations lors de l'échange vocal (Beebe *et al.* 1985; Jaffe *et al.* 2001).

Outre le possible apprentissage social de la temporalité des productions vocales, plusieurs exemples appuient l'idée selon laquelle il existe un apprentissage social progressif du contexte approprié d'utilisation des vocalisations au cours du développement ontogénétique

(e.g. Gouzoules et Gouzoules 1995,1989; Katsu *et al.* 2014; Laporte et Zuberbühler 2011). L'un des exemples les plus connus est celui de l'apprentissage de l'utilisation pertinente des vocalisations d'alarme spécifique à la détection d'un aigle chez le singe vervet (*Chlorocebus pygrythrus*). Au début de leur vie, les jeunes individus produisent cette vocalisation à l'encontre de tout item volant, notamment les feuilles d'arbres. Au cours de leur développement, les juvéniles émettent cette vocalisation uniquement envers les oiseaux, sans distinction. Enfin, les adultes produisent ces vocalisations d'alarme spécifique préférentiellement suite à la détection des rapaces, surtout l'aigle martial (*Polemaetus bellicosus* ; Seyfarth et Cheney 1986; Seyfarth *et al.* 1980). L'expérience et l'observation des adultes auraient un rôle primordial dans cet apprentissage. De plus, l'exposition répétée aux interactions entre certains membres du groupe peut également fournir des informations au jeune quant au comportement vocal à adopter pour interagir avec certains individus. C'est notamment le cas de l'utilisation des vocalisations « d'accueil » (*i.e. greeting calls*) qui sont produites à l'encontre d'un supérieur hiérarchique chez les chimpanzés (Laporte et Zuberbühler 2010). La production de cette vocalisation s'affine au cours du développement du jeune, passant des vocalisations non-dirigées socialement, à des vocalisations produites vers un grand nombre d'individus pour finir par n'être émise qu'en direction du supérieur hiérarchique une fois le stade sub-adulte atteint (Laporte et Zuberbühler 2011). Parfois, les adultes participent activement à l'apprentissage de l'utilisation de certaines vocalisations par les jeunes (Caro et Hauser 1992). C'est notamment le cas chez les tamarins : lors du transfert d'aliment de l'adulte vers le jeune, la production de vocalisations alimentaires par l'adulte, associée à ce transfert de nourriture, peut permettre au jeune d'apprendre les vocalisations appropriées au contexte alimentaire (Roush et Snowdon 2001). De même chez les marmosets, il est observé que les jeunes individus élevés au contact de leurs parents acquièrent plus précocement les structures vocales favorisant les échanges organisés (*i.e. vocalisation de contact*) avec les adultes que des juvéniles qui grandissent uniquement aux côtés d'individus du même âge (marmoset : Takahashi *et al.* 2017). Ce phénomène peut s'expliquer par le renforcement qu'induisent les parents en répondant aux vocalisations appropriées de leurs jeunes et en ignorant les vocalisations inappropriées (Chow *et al.* 2015; Takahashi *et al.* 2017).

Enfin, le jeune doit être en mesure de reconnaître l'identité de l'émetteur d'une production vocale pour envisager de lui répondre, car le choix des interlocuteurs lors d'un échange vocal n'est pas aléatoire ni opportuniste mais dépend de facteurs intrinsèques et/ou extrinsèques propres aux individus (**article 1**). Dès leurs plus jeunes âges, les macaques japonais semblent être en mesure de discriminer l'identité des émetteurs, dont celle de leur mère (Masataka 1985). Toutefois, il a été montré expérimentalement que les jeunes macaques, à



l'inverse des adultes, sont incapables de percevoir les phénomènes de convergences vocales qui se manifestent entre deux congénères pendant un échange vocal (Bouchet *et al.* 2017). Cette capacité de discrimination s'améliore avec le développement ontogénétique de l'individu (Bouchet *et al.* 2017), conjointement à la capacité de ces mêmes individus à converger vocalement avec leurs congénères au cours d'un échange vocal organisé (Masataka 2003; Sugiura et Masataka 1995; Sugiura 1998).

La maîtrise de la temporalité adéquate aux échanges vocaux, de l'utilisation des vocalisations appropriées au contexte et à l'interlocuteur semble être sous-tendue par un apprentissage social favorisé par des « tuteurs » sociaux servant de modèle au jeune, parallèlement à l'acquisition d'expérience sociale : l'exposition et la participation fréquente du juvénile à des échanges vocaux serait une condition préalable pour la pleine maîtrise des échanges vocaux organisés et les règles temporelles et sociales qui les régissent.

#### 2.2.2. Influence de la motivation sociale

Chez les gorilles des plaines de l'Ouest, nous avons pu remarquer que les individus les plus jeunes portaient une attention importante aux échanges vocaux impliquant un individu de leur classe d'âge (**article 2**). De même, les individus les plus âgés étaient plus attentifs vis-à-vis d'un échange vocal incluant des animaux adultes que des jeunes individus (**article 2**). Au regard de ces résultats, nous pouvons supposer que les gorilles possèdent une motivation particulière à tenir compte des interactions impliquant un membre du groupe qui fait partie de leur environnement social proche. Il est possible qu'à l'écoute de la vocalisation, les individus testés aient cherché à interagir avec l'émetteur, comme le suggèrent Lemasson et collaborateurs (2011a) qui rapportent des réponses comportementales similaires chez les mones de Campbell. Nous pouvons également imaginer que cette attention toute particulière des sujets découle du besoin des individus d'évaluer les relations existantes entre les autres membres du groupe, particulièrement celles impliquant des congénères avec lesquels ils interagissent fréquemment, afin de réagir en conséquence (Freeberg *et al.* 2012; Hinde 1976).

En suivant cette réflexion, nous pouvons supposer que l'intégration sociale du jeune au sein du groupe, et donc la construction d'un réseau social pour ce dernier, joue un rôle clé dans l'acquisition des règles « conversationnelles », et donc dans la maîtrise des échanges vocaux organisés. D'ailleurs, l'intégration sociale du jeune est corrélée avec le développement de sa maîtrise des règles « conversationnelles » (Eaton *et al.* 1985; Katsu *et al.* 2014; Lemasson *et al.* 2013a).

En effet, dans les premières années de la vie d'un jeune primate, les interactions sociales entre celui-ci et les autres membres du groupe sont rares. Le comportement du juvénile est encore fortement influencé par celui de sa mère (*e.g.* macaque rhesus : De Waal 1996). Au cours de son développement ontogénétique, le juvénile est amené à s'émanciper de sa mère et à diversifier ses relations sociales avec ses autres congénères (Eaton *et al.* 1985). L'intégration sociale du jeune dans le groupe peut faire naître chez lui le besoin de réguler les relations qu'il entretient avec ses congénères, mais aussi d'appréhender les relations existantes entre les autres membres de la communauté à laquelle il appartient (Freeberg *et al.* 2012; Hauser 1996; Owings et Morton 1998). Nous savons aujourd'hui que les échanges vocaux ne se manifestent pas aléatoirement au sein du groupe social, mais entre des interlocuteurs choisis en fonction de leurs liens affiliatifs (*e.g.* macaque japonais : Arlet *et al.* 2015; singe araignée : Briseño-Jaramillo *et al.* 2018), de leur âge (Lemasson *et al.* 2010, 2018) ou encore de leur statut hiérarchique (*e.g.* capucin à face blanche : Digweed *et al.* 2007). En ce sens, les échanges vocaux sont des outils cruciaux qui peuvent prendre une importance croissante pour l'individu, en lien avec son degré d'intégration sociale dans le groupe.

Par ailleurs, lors d'un échange vocal, il est possible que le manque de synchronisation vocale avec un congénère témoigne d'un manque d'attention ou d'expérience de l'interlocuteur (Henry *et al.* 2015). Ce défaut, qui soulignerait l'inexpérience de l'individu, pourrait lui desservir en le rendant peu fiable socialement auprès des autres membres de son groupe (Henry *et al.* 2015) et donc rendre plus difficile son intégration sociale. Par exemple, la production non appropriée des vocalisations d'alarme spécifique à la détection d'un aigle par les jeunes individus sont généralement ignorées par les autres membres de la communauté (Seyfarth *et al.* 1980). Il est possible que cette pression sociale supplémentaire favorise l'acquisition de la bonne maîtrise des règles conversationnelles au cours du développement du jeune.

En résumé, la maîtrise des échanges vocaux organisés semble être soumise à des processus de maturation des substrats cérébraux ainsi qu'à des phénomènes complexes d'apprentissage, d'expérience et de motivation sociale (*i.e.* intégration sociale). Des études complémentaires sont cruciales pour évaluer le rôle respectif de ces différents mécanismes au cours du développement ontogénétique. L'importance du respect des règles qui régissent les échanges vocaux et leur acquisition interroge intuitivement la fonction adaptative d'un tel comportement vocal.

### 3. Importance des règles et fonctions des échanges vocaux

#### 3.1. Importance du respect des règles

L'importance que témoignent les gorilles au respect des règles « conversationnelles » (**article 2**) est en accord avec les résultats expérimentaux obtenus chez les mones de Campbell (Lemasson *et al.* 2011), les macaques japonais (Bouchet *et al.* 2017) ou encore les bonobos (Levréro *et al.* 2017). Pour tenter de comprendre l'importance que semble avoir le respect des règles temporelles et/ou sociales pour les individus, plusieurs pistes de réflexions non-exclusives sont à explorer.

Lors de la production d'un signal acoustique, une information est transmise (volontairement ou non) et est perçue par un receveur (Seyfarth et Cheney 2003; Smith 1965). Ce signal acoustique peut encoder des informations quant à l'identité de l'émetteur comme son âge, son sexe, son rang social ou encore son appartenance à un groupe social (*e.g.* Miller, Mandel *et al.* 2010; Norcross et Newman 1993; Ramos-Fernández 2005; Salmi *et al.* 2014; Teixidor et Byrne 1999). Par conséquent, l'évitement de la superposition des signaux acoustiques par le respect du délai inter-vocalisations permettrait d'extraire sans ambiguïté l'information transmise par la vocalisation initiale (Seyfarth *et al.* 2010; Takahashi *et al.* 2013a). Des conclusions similaires sont énoncées pour expliquer, en partie, le respect du délai de silence séparant les productions vocales chez l'espèce humaine (Duncan 1972; Levinson 2016), malgré des différences sémantiques qui semblent indéniables.

Par ailleurs, l'organisation temporelle peut, en elle-même, être porteuse d'information pour les interlocuteurs. Chez l'humain, la coupure de parole est perçue de façon très négative (**article 1**, voir Goldberg 1990 et Kurtić *et al.* 2013 pour de rares cas de perception positive de la superposition des vocalisations chez l'humain). Elle est souvent attribuée à un fort manque de respect et de politesse (Calame-Griaule 1965; Ter Maat *et al.* 2010, voir **article 1**). Il en va de même pour certaines espèces d'oiseaux chez lesquelles le recouvrement de la vocalisation d'un congénère est considéré comme signal agressif fort (*e.g.* troglodyte mignon : Naguib et Kipper 2006) et mettant fin à l'interaction (*e.g.* Baker *et al.* 2012; Henry *et al.* 2015; Naguib et Kipper 2006; Naguib et Mennill 2010, voir *Introduction générale*, partie 3.1.3.). Si ce non-respect du délai inter-vocalisations est volontaire (Kurtić *et al.* 2013), il n'est pas exclu qu'il soit également un signal agonistique pour certaines espèces de primates non-humains.

Au-delà de l'importance du respect des règles pour les interlocuteurs eux-mêmes, les caractéristiques temporelles des échanges vocaux organisés (*i.e.* rupture ou non du délai inter-vocalisations), ainsi que les phénomènes de convergence vocale qui témoignent des affinités

sociales fortes (*e.g.* singe de Diane : Candiotti *et al.* 2012; mone de Campbell : Lemasson *et al.* 2003; marmoset : Snowdon et Elowson 1999) ou d'un rapport hiérarchique établi entre les interlocuteurs (macaque japonais : Lemasson *et al.* 2016), peuvent constituer une mine d'informations sociales essentielles pour l'audience, reflétant l'organisation sociale du groupe. Cela expliquerait l'attention particulière que portent les membres de l'audience aux respects des règles « conversationnelles ».

### 3.2. Fonctions des échanges vocaux organisés

Plusieurs fonctions sociales peuvent être attribués aux échanges vocaux au regard du type d'informations (*i.e.* identité de l'émetteur) qui transite lors de ces interactions. Tout d'abord, dans un environnement visuellement obstrué (*e.g.* couvert végétal dense), ces échanges peuvent être un moyen efficace de localiser les congénères (*e.g.* Candiotti *et al.* 2012). Cela facilite la cohésion socio-spatiale entre les membres du groupe (*e.g.* Harcourt *et al.* 1993; Ramos-Fernández 2005; Schamberg *et al.* 2016, 2017; Stewart et Harcourt 1994) et la coordination lors des déplacements (*e.g.* Sugiura 1993, 1998). Certains auteurs suggèrent également que les échanges « conversationnels » auraient pour fonction de réguler la compétition alimentaire par le maintien des distances inter-individuelles (Spehar et Di Fiore 2013). Ainsi, les échanges vocaux permettraient de temporiser les potentielles agressions et de favoriser la tolérance à l'approche d'un congénère (Kulahci *et al.* 2015; Seyfarth *et al.* 1994; Spehar et Di Fiore 2013). Des résultats obtenus chez les gorilles des montagnes vont dans ce sens en stipulant que la structure acoustique (*i.e.* formant bas) du *double-grunt* initiateur d'une interaction encode la volonté de l'émetteur d'entrer en compétition pour une ressource (Seyfarth *et al.* 1994). La structure acoustique du *double-grunt* émis en réponse (*i.e.* formant haut) exprimerait la soumission de l'interlocuteur en vue d'éviter un conflit coûteux pour les deux protagonistes (Seyfarth *et al.* 1994). Toutefois, les auteurs n'écartent par le fait que la vocalisation initiatrice de l'interaction (*i.e.* formant bas) témoignerait de la volonté de se déplacer vers son congénère, et que la réponse (*i.e.* formant haut) traduirait sa tolérance vis-à-vis de l'approche (Seyfarth *et al.* 1994).

Cependant, un nombre croissant d'éléments, aussi bien comportementaux que physiologiques, soutiennent le rôle consolidateur de liens sociaux des échanges vocaux comme le propose Dunbar (1996).

Le réseau de toilettage physique peut être retracé au regard du réseau d'échanges vocaux chez certaines espèces de primates non-humains (*e.g.* macaque japonais : Arlet *et al.* 2015;

tamarin : Jordan *et al.* 2004; maki catta *Lemur catta* : Kulahci *et al.* 2015; mone de Campbell : Lemasson *et al.* 2006; marmoset : Snowdon et Cleveland 1984; singe écureuil *Saimiri sciureus* : Soltis *et al.* 2002), mais aussi chez quelques espèces de mammifères non-primates (*e.g.* orque : Miller *et al.* 2004; éléphant : Soltis *et al.* 2005). En ce sens, l'étude comparative que nous avons menée chez les grands singes (**article 4**) corrobore cette théorie. Chez les espèces pour lesquelles nous retrouvons des échanges vocaux organisés, le réseau affiliatif reflète le réseau que constituent les échanges vocaux (Lemasson *et al.* 2018; Levréro *et al.* 2019, **article 4**). L'aspect particulièrement affiliatif de ces interactions vocales se manifeste également par les phénomènes de convergence vocale, favorisant l'intégration et renforçant les affinités entre interlocuteurs (Lemasson et Hausberger 2004; Sewall *et al.* 2016), et la réponse vocale rapide envers les interlocuteurs affiliés (*e.g.* Biben *et al.* 1986; Masataka et Biben 1987).

Les échanges vocaux pourraient également avoir pour bénéfice de réduire le stress des interlocuteurs. En effet, une réduction notable du taux de cortisol urinaire est rapportée chez les individus suite à l'écoute de vocalisations de leurs congénères (marmoset : Owren *et al.* 2011; Rukstalis et French 2005).

Un parallèle intéressant peut-être tracé entre la fonction probable des échanges vocaux chez les primates non-humains et celle répertoriée chez l'espèce humaine. Au stade pré-verbal, les enfants humains maîtrisent précocement le tour de parole et s'engagent fréquemment dans des échanges vocaux organisés temporellement avec leurs parents ou tuteurs (Gratier *et al.* 2016; Hilbrink *et al.* 2015; Jaffe *et al.* 2001). Pourtant, quand les enfants commencent à interagir avec leurs semblables en utilisant le langage, ils éprouvent des difficultés à alterner avec fluidité leurs vocalisations avec celles de leurs interlocuteurs (Hilbrink *et al.* 2015). Cette difficulté se manifeste par un allongement du délai inter-vocalisations (Hilbrink *et al.* 2015). Selon certains auteurs, la perte de synchronisation provient de la difficulté que les jeunes enfants ont à formuler une réponse en lien avec leur maîtrise approximative de la syntaxe et de la sémantique (Casillas *et al.* 2016; Levinson 2016; Lindsay *et al.* 2019). Nous pouvons alors émettre l'hypothèse suivante : lors du stade pré-verbal, la synchronisation vocale entre le jeune et l'adulte aurait pour fonction de maintenir et de renforcer la cohésion sociale entre les individus (Reddish *et al.* 2014). Une fois le stade verbal atteint, les conversations rempliraient une seconde fonction attenante au contenu sémantique du message qui est transmis (Dunbar *et al.* 1997), allant des informations relevant des relations et des expériences personnelles, à la planification d'activités futures (sociales ou non), en passant par des éléments extérieurs ancrés dans le présent (*e.g.* divertissements, politique, religion, travail ou informations techniques). En suivant cette réflexion, nous pouvons supposer que les échanges vocaux des primates non-

humains remplissent des fonctions similaires à celles des interactions des enfants pré-verbaux. Néanmoins, une différence qualitative majeure semble se dessiner entre les primates non-humains et les humains « verbaux » quant aux fonctions de ces « conversations ». Mis à part le contenu informatif individu-centré encodé par les vocalisations utilisées (*e.g.* identité de l'émetteur, son état émotionnel et ses intentions propres) permettant de maintenir la cohésion socio-spatiale, nous n'avons aucune preuve que les échanges vocaux des primates non-humains facilitent la communication sur un élément de l'environnement externe aux individus.

En résumé, les échanges vocaux transmettraient des informations relatives à l'identité des interlocuteurs, leurs intentions et la qualité des liens sociaux existants entre eux. Cette fonction informative serait essentielle à la cohésion socio-spatiale des membres du groupe. Toutefois, bien que des similitudes fonctionnelles puissent être suggérées entre les échanges vocaux des primates non-humains et ceux des enfants humains pré-verbaux, des différences sémantiques majeures sont ensuite à souligner entre les interactions vocales simiennes et les conversations humaines. Le contenu sémantique des échanges vocaux organisés doit encore être exploré de façon plus systématique.

#### 4. Emergence des échanges vocaux

##### 4.1. Des facteurs sociaux essentiels à la mise en place des échanges vocaux ?

Dans son ouvrage *"Grooming, gossip et the evolution of language"* paru en 1996, Dunbar suggère que la capacité à « converser » a joué un rôle primordial au cours de l'évolution des primates. Selon cet auteur, la vie sociale des primates a gagné en complexité suite à une augmentation de la taille de leurs groupes sociaux, et par conséquent du nombre de partenaires affiliés. Dans ce contexte, les primates n'étaient plus en mesure de témoigner, par le toilettage social, leurs affinités à l'ensemble de leurs congénères : le toilettage social manuel se serait donc mué en toilettage social à distance prenant la forme d'échanges vocaux temporellement et socialement structurés. Ainsi, la complexification des groupes sociaux aurait mené à une complexification de la communication vocale chez les primates (Dunbar 1993, 1996, 1998). De cette réflexion, nous supposons que les échanges vocaux sont la conséquence d'une vie sociale en groupe de taille importante, impliquant de nombreuses interactions affiliatives qui se manifestent par un taux élevé de toilettage physique.

Nos investigations confirment que la vie sociale est une condition indispensable pour voir émerger des échanges vocaux chez une espèce donnée. Cet aspect est illustré par l'absence

presque totale d'interactions vocales (*i.e.* duo, *chorus* et échanges vocaux) reportée chez les orangs-outans de Bornéo (**article 4**), connus pour leur faible propension à la socialité (*e.g.* Delgado et van Schaik 2000; Galdikas 1985a; Knott *et al.* 2008; Knott 1999)

Cependant, nos observations ne révèlent pas de corrélation entre l'effectif du groupe, le taux de toilettage social et le taux d'échanges vocaux, contrairement aux prédictions de Dunbar (1996). En effet, des échanges vocaux organisés sont reportés aussi bien chez les bonobos qui évoluent dans des communautés regroupant 30 à 80 individus, que chez les gorilles des plaines de l'Ouest qui constituent des groupes sociaux d'une dizaine d'individus (**article 4**). De plus, le taux de toilettage physique chez ces deux espèces est inférieur à celui exprimé par les chimpanzés qui ne produisent que très rarement des échanges vocaux basés sur le tour de parole et qui vivent dans des communautés de taille comparable à celle de leur espèce « sœur » (**article 3** et **article 4**).

Toutefois, nous apportons des précisions supplémentaires quant aux facteurs pouvant favoriser l'émergence d'échanges vocaux organisés chez une espèce donnée. Chez les chimpanzés, le réseau affiliatif de chaque individu est très sélectif et se restreint à quelques partenaires de rang hiérarchique équivalent, comme en témoigne leur comportement de toilettage social (Girard-Buttoz *et al.* 2020). Leur société est également régie par un degré de compétition élevée et une territorialité marquée (*e.g.* Müller et Mitani 2005; Watts et Mitani 2001). À l'inverse, les bonobos développent un réseau affiliatif important et diversifié (Girard-Buttoz *et al.* 2020) et sont connus pour leur fort degré de tolérance vis-à-vis des membres de leur communauté et des communautés voisines (*e.g.* Boesch *et al.* 2002; Gómez *et al.* 2016; Wilson *et al.* 2014). De même, un certain degré de tolérance ainsi que des liens affiliatifs stables et durables s'établissent entre les femelles gorilles des plaines de l'Ouest et le mâle adulte à la tête du groupe (Cooksey *et al.* 2020; Forcina *et al.* 2019; Watts 2002). La tolérance intra-groupe élevée est une caractéristique commune aux bonobos et aux gorilles des plaines de l'Ouest qui pourrait sous-tendre la présence d'échanges vocaux chez ces deux espèces, et son absence chez les chimpanzés plus compétitifs (**article 4**). Chez ces derniers, les interactions vocales sont majoritairement exprimées sous la forme de *chorus* de vocalisations de longue distance (*i.e.* *pant-hoot*) qui se manifestent entre interlocuteurs ayant des affinités sociales (Fedurek *et al.* 2014, 2013ab; Mitani et Gros-Louis 1998). Dans la société hautement compétitive et territoriale des chimpanzés, ce type d'interactions vocales permettent, entre autres, d'intimider de potentiels rivaux et ainsi d'éviter des conflits coûteux pour les individus (Wilson *et al.* 2001).

Il reste difficile d'établir une relation entre les critères classiquement utilisés pour mesurer la complexité sociale (*e.g.* taille du groupe) et ceux permettant d'appréhender la complexité vocale (*e.g.* taille du répertoire vocale). Toutefois, la nature de la société (diversité et valence des liens sociaux) et le type d'interactions vocales (absence d'interactions, *chorus* et échanges vocaux organisés) semblent corrélés. En ce sens, l'étude des interactions vocales pourrait constituer un nouvel outil prometteur pour appréhender la qualité des liens sociaux.

A la vue de nos résultats, nous pouvons également supposer que la tolérance intra-groupe est un indicateur de la complexité sociale pouvant être liée à une certaine complexité communicative. Cette réflexion est soutenue par de récentes observations menées chez quatre espèces de macaques (*Macaca fuscata*, *Macaca mulatta*, *Macaca nigra* et *Macaca tonkeana*) qui prouvent que les espèces les plus tolérantes présentent un niveau de diversité et de flexibilité vocale plus élevée que les espèces peu tolérantes (*e.g.* interactions agonistiques ou affiliatives, Rebout *et al.* 2020). Les auteurs suggèrent qu'une forte tolérance, en donnant moins de poids au statut social, élargit l'éventail des issues possibles à leurs interactions sociales. A l'inverse, une intolérance marquée, qui donne plus de poids au statut social, rend les débouchés des interactions bien plus prévisibles (Duboscq *et al.* 2017; Thierry 2007). Etant donné que l'incertitude semble être l'une des caractéristiques majeures des systèmes complexes (McDaniel et Driebe 2005; Schuster 2016), il est possible qu'en parvenant à évaluer objectivement le degré de tolérance d'une espèce, nous puissions lui attribuer un degré de complexité sociale. En ce sens, Rebout et collaborateurs (soumis) ont développé un modèle prometteur permettant d'appréhender le degré d'incertitude (*i.e.* qui témoigne du degré de tolérance) d'un système social en s'inspirant du modèle proposé par Shannon (1948) dans sa *Théorie de l'information*. Ainsi, les auteurs décomposent la complexité en trois dimensions que sont la diversité (*e.g.* l'incertitude est plus grande dans les groupes qui possèdent une composition équilibrée d'individus de chaque catégorie d'âge et de sexe), la flexibilité (*e.g.* flexibilité fonctionnelle : plus un comportement est produit dans des contextes différents, et plus l'incertitude est élevée) et la combinabilité (*e.g.* lors d'un conflit, plus les adversaires émettent un nombre équivalent d'agression et plus l'incertitude est élevée). En quantifiant ces trois dimensions à l'aide de la formule de Shannon (1948), les auteurs peuvent obtenir un indice de complexité tripartite permettant de juger objectivement de la complexité d'un système et de pouvoir le comparer à un autre. Rebout *et al.* (2020) suggèrent également d'appliquer cette méthodologie pour appréhender la complexité vocale.

Aujourd'hui, il est essentiel de multiplier les études comparatives ciblées sur des espèces pertinemment choisies pour identifier les facteurs sociaux multiples et subtiles sous-jacent à la



complexité sociale, et pouvant être à l'origine de la complexité des systèmes communicatifs. Par exemple, en sélectionnant des espèces qui n'émettent pas d'échanges vocaux organisés (ou qui les utilisent mais avec des règles différentes de celles connues jusque-là), nous pourrions identifier les facteurs qui conditionnent l'émission de ces interactions vocales « conversationnelles ».

#### 4.2. L'apparition des échanges vocaux : convergence évolutive ou caractère ancestral ?

Les nombreuses découvertes d'échanges vocaux temporellement et socialement organisé sur toutes les branches de la phylogénie des primates non-Hominidés (*e.g.* Prosimiens : maki catta : Kulahci *et al.* 2015 ; Plathyrriniens : singe écurueil : Biben *et al.* 1986; singe araignée : Briseño-Jaramillo *et al.* 2018; capucin : Digweed *et al.* 2007; marmoset pygmé : Snowdon et Cleveland 1984; marmoset commun : Takahashi *et al.* 2013a ; Cercopitecoïdes : singe Diane : Candiotti *et al.* 2012; mone de Campbell : Lemasson *et al.* 2010; macaque japonais : Sugiura 1993; voir **article 1**), couplé à la mise en évidence d'échanges vocaux chez certaines espèces de grands singes (*e.g.* gorille des plaines de l'Ouest : Lemasson *et al.* 2018; bonobo : Levréro *et al.* 2019; **articles 4**), tendent à rendre plausible l'idée selon laquelle l'origine de nos conversations est profondément ancrée dans la lignée des primates. Ce comportement vocal est probablement encore plus ancien, comme le suggère les études qui prouvent l'existence d'échanges vocaux temporellement organisés chez des taxons plus éloignés de mammifères (*e.g.* orque : Miller *et al.* 2004; cachalot : Schulz *et al.* 2008; éléphant : Soltis *et al.* 2005; rat-taupo nu : Yosida *et al.* 2007; Yosida et Okanoya 2009), et même d'oiseaux (*e.g.* etourneau : Henry *et al.* 2015; corbeau à gros bec : Kondo *et al.* 2010).

L'absence d'interactions vocales « conversationnelles » chez les orangs-outans de Bornéo et les chimpanzés pourrait résulter de pressions de sélection différentes, conduisant à l'apparition de traits évolutifs dérivés. Dans le cas des orangs-outans de Bornéo, la vie majoritairement solitaire ne leur donnerait alors pas l'occasion, ni le besoin, de produire ce type de comportements vocaux (**article 4**). Concernant les chimpanzés, l'absence d'échanges vocaux est plus délicate à justifier, mais des pistes de réflexions se dessinent au regard de la compétitivité accrue de leur système social, de la sélectivité de leur réseau social affiliatif et de l'efficacité de leur comportement de toilettage social, notamment dû au nombre limité de partenaires sociaux (**article 4**). L'absence de preuves d'échanges vocaux organisés temporellement et socialement chez ces deux espèces ne signifie pas pour autant qu'ils n'engagent pas d'interactions « conversationnelles » par le biais d'autres modalités

communicatives. Les résultats prometteurs de Fröhlich et collaborateurs (2016) qui révèlent l'existence d'interactions multi-modales organisées sur le « tour de parole » régissant la communication mère-enfant chez les chimpanzés en sont une preuve. Les capacités communicatives des chimpanzés et des orangs-outans aux travers des modalités non-vocales ne sont donc pas à négliger et doivent être abordées méticuleusement dans le futur (*e.g.* chimpanzé : Graham *et al.* 2018; Hobaiter et Byrne 2011; orang-outan : Liebal *et al.* 2006).

Par ailleurs, certains contextes n'ont pas été répertoriés dans nos études. Nous pouvons notamment penser aux comportements de chasses collaboratives des chimpanzés en milieu naturel (*e.g.* Boesch 1994; Mitani et Watts 2001). En effet, chez l'humain, l'idée selon laquelle le langage s'est développé dans un contexte de chasse, pour communiquer sur la localisation et la planification de l'action, a persisté pendant plusieurs années (Dunbar 1993). Toutefois, la chasse organisée chez les chimpanzés est peu fréquente et a été reportée principalement dans une communauté (communauté de Taï, Côte d'Ivoire : 63% des événements de chasse étaient collaboratifs, contre seulement 7% dans la communauté de Gombe, Tanzanie : Boesch et Boesch 1989; Busse 1977; Goodall 1986). En outre, dans un contexte de chasse, il est peu probable que des vocalisations soient produites, l'intérêt étant de ne pas être détecté par la proie (Graw et Manser 2007; Stone et Trost 1991).

Nos résultats chez les grands singes viennent compléter nos connaissances sur l'utilisation des échanges vocaux, définis par des règles temporelles précises, chez les primates sociaux en général. Pourtant, il n'est en aucun cas possible d'affirmer que cette propriété langagière est issue d'un héritage commun aux primates plutôt que d'une convergence évolutive médiée par les besoins de la vie sociale. En effet, d'autres espèces hautement sociales, et phylogénétiquement éloignées des primates, utilisent également des échanges vocaux organisés.

Par ailleurs, malgré des similitudes indéniables au niveau de l'organisation temporelle entre les conversations humaines et les échanges vocaux organisés des primates non-humains, des différences majeures demeurent. En plus des différences fonctionnelles que nous avons abordées précédemment (voir partie 3.2.), des différences structurelles se manifestent également. Les échanges vocaux décrits chez les primates non-humains se résument généralement à un seul tour de parole bref entre l'émetteur et son interlocuteur (Lemasson *et al.* 2010, 2018; Levréro *et al.* 2019). Seul les marmosets font exception puisque, à l'image de l'humain (*e.g.* Hayashi 2013; Sacks *et al.* 1974), leurs échanges vocaux se composent de nombreuses alternances de prises de paroles (Takahashi *et al.* 2013a). Certains auteurs

attribuent cette similarité entre humains et marmosets au développement d'une stratégie de reproduction coopérative impliquant une tendance importante à la pro-socialité telle que des liens sociaux solides, une grande tolérance au niveau du groupe, une résolution coopérative des problèmes et un partage actif de la nourriture (Burkart *et al.* 2009; Burkart et van Schaik 2010; Snowdon et Cronin 2007; Takahashi *et al.* 2013a). Ceci renforce l'idée que les besoins sociaux sont le moteur principal pour l'émergence et la complexification de l'usage des échanges vocaux organisés.

Les possibles dissemblances fonctionnelles (*i.e.* d'ordre qualitative) et structurelles (*i.e.* d'ordre quantitative) que nous mettons en lumière entre les conversations humaines et les échanges vocaux organisés des primates non-humains ouvrent un vaste champ d'investigation.

En résumé, en dépit des résultats prometteurs que nous avons obtenus, le scénario évolutif des conversations humaines reste difficile à décrire en s'appuyant sur les plus proches parents du genre *Homo* : les grands singes. Les aptitudes à interagir sous la forme d'échanges vocaux organisés résultent probablement de multiples pressions évolutives. Les primates humains comme non-humains partagent tout de même certaines règles temporelles et sociales qui régissent leurs échanges vocaux pouvant être attribuées à des contraintes similaires. Il est possible que les chimpanzés et les orangs-outans de Bornéo aient été soumis à des pressions sélectives différentes, respectivement une forte compétition et une vie majoritairement solitaire, menant à l'apparition de caractères évolutifs dérivés. De plus amples investigations sont nécessaires pour évaluer la propension de ces espèces à « converser » à une échelle multi-modale. Enfin, nos conversations se distinguent des échanges vocaux organisés des primates non-humains sur les plans fonctionnels (*i.e.* contenu sémantique) et structurels (*i.e.* nombre de tours de parole). Toutefois, la capacité à interagir vocalement avec ses congénères, même si elle est limitée à de courts échanges, pourrait avoir accompagné la complexification de la vie sociale et sous-tendre une évolution divergente de la communication vocale des primates (Dunbar 1996).

## 5. Perspectives majeures

Le travail fourni tout au long de ce projet de recherche doctorale doit être poursuivi et approfondi sur plusieurs aspects que sont la mise en place d'observations et d'expérimentations supplémentaires, des précisions quant aux protocoles appliqués et la suggestion de nouvelles questions de recherches ayant trait aux échanges vocaux.

### 5.1. Des précisions à apporter

Tout d'abord, l'effectif de nos échantillons a pu être limitant à plusieurs reprises. C'est notamment le cas lors des expérimentations que nous avons menées sur les gorilles des plaines de l'Ouest (**article 2**). Une plus grande représentation des différentes classes d'âge permettrait de renforcer nos découvertes quant à la pertinence des règles temporelles et sociales pour cette espèce ou, à l'inverse, d'apporter de nouveaux éléments de discussion. Aussi, nos résultats concernant l'attention portée aux règles en fonction de l'âge des individus pourraient être complétés et précisés évaluant les réponses comportementales des individus jeunes proches de leur mère. Toutefois, il faudra tenir compte de l'attention sociale conjointe qui peut influencer la réponse attentionnelle des jeunes et qui peut représenter une des limites majeures de la situation expérimentale (Emery 2000).

Par ailleurs, dans le cas des observations menées sur les chimpanzés et les orangs-outans de Bornéo, une plus grande représentation des différentes classes d'âge et de sexe aurait pu être bénéfique pour la description de l'organisation temporelle des productions vocales (**article 3** et **article 4**).

Les protocoles observationnels et expérimentaux que nous avons utilisés peuvent également être améliorés et précisés (**article 3** et **article 4**). Comme nous l'avons vu précédemment, il est probable que les chimpanzés, mais aussi les autres espèces de grands singes, s'impliquent dans des échanges « conversationnels » via d'autres modalités que les seules vocalisations (voir partie 4.2.). Ainsi, il serait pertinent d'avoir un approche multi-modale chez ces espèces pour évaluer leur propension à « converser », et ce, en tenant compte de leur répertoire communicatif non-vocal.

Les résultats que nous avons obtenus suite à l'expérimentation basée sur le protocole de violation des attentes chez les gorilles des plaines de l'Ouest (**article 2**) soulèvent de nouvelles questions qui nécessitent des protocoles expérimentaux adaptés. La moindre attention que portent les jeunes individus aux échanges vocaux impliquant des adultes doit par exemple être considérée. Ainsi, nous pouvons penser que la réponse attentionnelle des jeunes à l'écoute d'échanges simulés impliquant des adultes peut être influencée par l'âge des interlocuteurs plus que par le respect de la règle temporelle. Pour trancher sur cette question, la réaction des jeunes individus peut être intéressante à comparer entre deux situations : (i) une coupure de parole impliquant deux adultes, et (ii) une coupure de parole impliquant un jeune initiateur de l'échange et un adulte répondant. Par ailleurs, l'implication systématique du mâle adulte en tant que répondant dans nos échanges vocaux simulés (**article 2**), bien que calquée sur nos propres

observations et sur la littérature (Lemasson *et al.* 2018), peut aussi influencer la réponse attentionnelle de l'audience quand ce dernier brise une règle temporelle et/ou sociale. Etant donné sa position privilégiée au sein du groupe (*e.g.* Harcourt et Stewart 2007; Margulis *et al.* 2003), la violation d'une règle conversationnelle de sa part pourrait être particulièrement tolérée par l'audience. Ainsi, le fait de réitérer ce type d'expérimentation en simulant une violation des règles conversationnelles par un autre individu que le mâle pourrait révéler l'importance capitale de ce dernier lors de ces interactions vocales, et la tolérance que témoignent les individus du groupe envers le mâle dominant. Toutefois, la multiplication des repasses de vocalisations est à temporiser pour éviter autant que possible l'habituation des sujets testés, ainsi que le potentiel stress occasionné par la mise en place de l'expérimentation.

Les résultats présentés dans ce manuscrit devront également être complétés par des analyses menées chez d'autres populations de grands singes, à la fois en captivité et en milieu naturel, afin d'évaluer l'effet de l'environnement captif sur les caractéristiques temporelles et sociales des règles « conversationnelles ». Toutefois, bien que l'étude des grands-singes apparaisse comme une étape nécessaire, cela ne représente qu'un seul modèle pour chaque intensité de vie sociale. Au regard de la diversité des espèces, provenant de clades distincts, qui utilisent des échanges vocaux organisés, ainsi que d'autres types d'interactions vocales (*i.e.* duo et *chorus*), l'évolution de la capacité à interagir vocalement a probablement été foisonnante au sein du règne animal et marquée par de multiples épisodes de divergence et de convergence. Il est important maintenant de ne pas se restreindre au clade des primates, et de questionner les facteurs sociaux qui peuvent favoriser l'émergence des différents types d'interactions vocales chez d'autres branches de l'arbre phylogénétique en adoptant une approche comparative sur des espèces pertinemment choisies. Par exemple, certaines espèces de cétacés sont connues pour produire des échanges vocaux temporellement organisés sur le tour de parole, c'est le cas du grand dauphin (Nakahara et Miyazaki 2011), alors que d'autres favorisent les productions vocales superposées, comme les cachalots (Schulz *et al.* 2008). Plusieurs taxons se porteraient particulièrement bien à ce type d'investigation, tels que les cinq espèces de mangoustes (*Suricata suricatta*, *Mungos mungo*, *Helogale parvula*, *Cynictis penicillata* et *Galerella sanguinea*) étudiées par Manser et collaborateurs (2014), qui évoluent dans des systèmes sociaux distincts mais qui restent relativement proche phylogénétiquement et avec une écologie comparable. En menant de tels travaux permettant de comparer des espèces n'utilisant pas d'échanges vocaux organisés, à l'instar des orangs-outans et des chimpanzés, à des espèces en produisant fréquemment, comme les bonobos et les gorilles des plaines de l'Ouest, nous

pourrions en apprendre d'avantage quant aux pressions sociales favorisant la mise en place de ce type d'interactions vocales « conversationnelles ».

Pour approfondir la compréhension des mécanismes neuronaux sous-tendant à la perception et au contrôle des productions vocales lors des échanges, des méthodes de neuro-imageries cérébrales non-invasives pourraient être mises en place, à l'image des TEP-IRM comme exprimé précédemment. En plus de préciser les substrats cérébraux impliqués dans un contexte d'interaction, une approche longitudinale chez des espèces à développement rapide (*e.g.* marmoset) permettrait éventuellement de visualiser l'implication de structures neuronales différentes en fonction de l'âge des individus, et/ou la maturation des systèmes déjà présents.

Enfin, à la vue des nombreuses études menées sur les échanges vocaux organisés des primates non-humains et la flexibilité de production dont ils semblent faire preuve, certaines questions de recherche peuvent être soulevées. C'est notamment le cas de l'aspect intentionnel de ces comportements vocaux qui, à notre connaissance, ne semble pas avoir été abordé dans la littérature.

## 5.2. Les échanges vocaux organisés : fruit d'une intentionnalité sous-jacente ou simple contagion émotionnelle ?

Selon Dennett (1983), l'intentionnalité dont fait preuve un émetteur au moment de la production d'un signal communicatif peut être classée selon cinq degrés. Trois d'entre eux sont présentés ici : (i) le degré zéro correspond à un signal communicatif non-intentionnel produit spontanément à la présentation d'un stimulus (*e.g.* vocalisations d'alarme chez les singes vervets); (ii) le premier degré d'intentionnalité équivaut à l'émission d'un signal volontairement produit par un émetteur à l'attention d'un destinataire dans le but de modifier le comportement de ce dernier (Hobaiter et Byrne 2013; Tomasello *et al.* 1985), l'émetteur adaptant sa stratégie pour atteindre son objectif (Roeder et Gosset 2001); (iii) le second degré d'intentionnalité est proposé quand l'émetteur agit avec l'objectif de modifier l'état mental du receveur, impliquant de ce fait une attribution d'un état mental au receveur par l'émetteur (Graham *et al.* 2019).

La recherche de l'intentionnalité dans la communication chez les animaux non-humains est majoritairement axée sur le premier niveau d'intentionnalité (Graham *et al.* 2019; Townsend *et al.* 2017). Selon Townsend et collaborateurs (2017), le premier degré d'intentionnalité se manifeste quand le signal communicatif produit est socialement dirigé (*i.e.* attention portée vers

l'audience, sensible à la présence et à la composition de celle-ci), est émis en vue d'atteindre un but précis (*i.e.* persistance et /ou élaboration tant que l'objectif n'est pas atteint), et qu'il modifie le comportement du receveur (*i.e.* se traduit par la réponse comportementale de celui-ci).

Plusieurs observations (**article 1**) peuvent laisser penser que les échanges vocaux organisés des primates non-humains sont intentionnels au regard des marqueurs comportementaux proposés par Townsend et collaborateurs (2017). Par exemple, en ce qui concerne l'individu initiateur de l'interaction chez les macaques Japonais, Koda (2004) parvient à mettre en évidence le fait qu'en l'absence d'une réponse rapide à sa première vocalisation, l'individu initiateur est enclin à produire à nouveau le même type de vocalisation consécutivement. De même, la vocalisation répétée est souvent plus longue et voit sa structure acoustique modifiée avec une fréquence fondamentale plus haute que la vocalisation initiale (Koda 2004). Cette vocalisation modifiée induit plus fréquemment une réponse vocale de la part des congénères (Koda 2004). Il semble donc que l'initiateur soit en mesure de produire de façon répétée (*i.e.* persistance) et d'élaborer son signal en vue d'atteindre un but précis, tout en modifiant le comportement de l'individu receveur qui se traduit par l'émission de la réponse vocale. Une fois la réponse obtenue, l'initiateur met fin à ses vocalisations consécutives, témoignant que son but est atteint (Koda 2004).

En revanche, il n'existe pas, pour l'heure, de preuves que cette vocalisation initiatrice de l'échange soit dirigée vers une audience particulière. Des pistes de réflexions à ce sujet se trouvent peut-être dans la réponse vocale de l'interlocuteur. Certains éléments peuvent laisser penser que la réponse vocale est socialement dirigée vers l'initiateur. Par exemple, l'interlocuteur semble choisir de répondre à une vocalisation en fonction des caractéristiques de l'émetteur telles que son âge (*e.g.* Lemasson *et al.* 2010), son statut hiérarchique (*e.g.* Digweed *et al.* 2007), ou encore le degré d'affinité existant entre eux (*e.g.* Arlet *et al.* 2015). De même, chez certaines espèces, l'interlocuteur converge vocalement avec l'individu initiateur au moment de la réponse (*e.g.* Lemasson *et al.* 2016), et ce délai de réponse peut également varier en fonction de l'identité de l'initiateur (*e.g.* Masataka et Biben 1987). Bien que ces éléments semblent suggérer un certain degré d'intentionnalité sous-jacent à la réponse vocale, aucune preuve ne permet d'exclure le fait que cette réponse est le résultat d'une contagion émotionnelle induite par la vocalisation initiatrice de l'échange (Graham *et al.* 2019). Pour contrôler l'aspect émotionnel de la communication dans ce contexte, des outils de mesure de l'état physiologique de l'émetteur (*e.g.* température corporelle, rythme cardiaque) au moment de la production vocale peuvent être proposés (Graham *et al.* 2019). Cependant, des

innovations sont encore essentielles pour développer des outils de mesure en adéquation avec l'observation des animaux, permettant des relevés en temps réel sur des individus mobiles et sans influencer sur leurs interactions sociales (Graham *et al.* 2019).

La fonction essentiellement informative des échanges vocaux peut représenter l'une des limites majeures à l'étude de l'intentionnalité de ce comportement. Un acte communicatif informatif, à l'inverse d'une communication impérative, n'aboutit pas nécessairement à une modification du comportement de l'individu cible. Cet aspect complexifie la validation des marqueurs comportementaux permettant de distinguer l'intentionnalité de premier degré et de degré zéro (Graham *et al.* 2019). Toutefois, l'information transmise lors de l'échange vocal peut avoir pour but de modifier l'état mental du receveur, propriété qui est associée à une communication intentionnelle de second degré. La problématique réside alors dans le fait de prouver qu'à l'image de l'humain (Crockford *et al.* 2012; Dennett 1983), le fait d'échanger vocalement est réalisé avec l'intention de modifier l'état mental des récepteurs (Graham *et al.* 2019). En ce sens, l'objectif peut être de tester la capacité des individus à attribuer un état mental à leurs receveurs par le biais d'expériences développées dans le but d'évaluer l'intentionnalité de second degré (Graham *et al.* 2019). Sa démonstration, bien qu'applicable à un nombre restreint d'espèces (Liebal *et al.* 2014a; Townsend *et al.* 2017), permettrait de contourner l'utilisation des marqueurs comportementaux dont la validité semble limitée (Graham *et al.* 2019).

## 6. Conclusion

Au terme de cette thèse, les études expérimentales et observationnelles que nous avons menées chez les grands singes apportent des éléments supplémentaires de réflexions concernant le développement des règles « conversationnelles » au cours de l'ontogénie, les causes proximales et ultimes des échanges vocaux ainsi que sur les facteurs sociaux à l'origine de ce comportement vocal, universellement répandu chez l'espèce humaine.

Malgré le fait que ce comportement vocal se développe rapidement au cours u développement ontogénétique, la maîtrise des règles « conversationnelles » n'est pas innée. La perception et l'utilisation adéquate des vocalisations produites lors de ces interactions pourraient être sous-tendues par des processus croisés de maturation des substrats neuronaux et des mécanismes complexes d'apprentissage, d'expérience et de motivation sociale. Cette acquisition est corrélée avec l'intégration sociale de l'individu au sein de sa communauté.



Ces règles « conversationnelles » pourraient avoir pour finalité principale de limiter la dégradation du contenu informatif porté par les vocalisations impliquées dans l'interaction. Mais il n'est pas exclu que le respect même de ces règles constitue une information sociale en lui-même. Par ailleurs, parmi les diverses fonctions qui peuvent être attribuées aux échanges vocaux, nos résultats soutiennent l'idée selon laquelle ce comportement est avant tout utile pour établir, maintenir et renforcer les liens sociaux affiliatifs entre les interlocuteurs.

Au regard des convergences et divergences entre les différentes espèces étudiées, nos résultats permettent de proposer des facteurs sociaux favorisant l'émergence des échanges vocaux organisés. Ainsi, une vie sociale riche en interactions sociales, un réseau important de partenaires affiliés et un niveau de tolérance élevé semblent être des conditions favorables à la mise en place de ce comportement vocal. Nos résultats supportent l'idée d'une complexité vocale socialement motivée.

Enfin, la description d'échanges vocaux organisés chez les grands singes permet de compléter l'arbre phylogénétique de ce comportement vocal, révélant sa relative continuité au sein de la lignée primates, incluant l'humain. Cependant, il reste impossible de dessiner avec certitude le scénario évolutif de nos conversations et des règles qui les régulent. Des investigations supplémentaires sont cruciales en vue de cibler les multiples et subtils facteurs sociaux qui sous-tendent la complexité sociale et qui peuvent être à l'origine de l'évolution de la communication vocale.

Les présents résultats pourront être complétés par des approches comparatives menées sur des espèces pertinemment choisies avec divers systèmes sociaux, aussi bien de primates non-humains que d'animaux non-primates. Les résultats ainsi obtenus pourront nous informer quant aux causes proximales et ultimes de ce comportement interactionnel. Aussi, l'intentionnalité qui sous-tend les échanges vocaux ouvre un champ d'étude vaste est rarement exploré.





# Chapitre VII

## Références bibliographiques



© Illustration de Thierry Pognault



- Aboitiz, F., & Garcia, R. (1997). The evolutionary origin of the language areas in the human brain. A neuroanatomical perspective. *Brain Research Reviews*, 25, 381–396.
- Adami, C. (2002). What is complexity? *BioEssays*, 24(12), 1085–1094. <https://doi.org/10.1002/bies.10192>
- Aitken, P. G. (1981). Cortical control of conditioned and spontaneous vocal behavior in rhesus monkeys. *Brain and Language*, 13(1), 171–184. [https://doi.org/10.1016/0093-934X\(81\)90137-1](https://doi.org/10.1016/0093-934X(81)90137-1)
- Altmann, J. (1974). Observational Study of Behavior: Sampling Methods. *Behaviour*, 49(4), 227–267.
- Ancrenaz, M., Gumal, M., Marshall, A. J., Meijaard, E., Wich, S. A., & Husson, S. (2018). *Pongo pygmaeus* (errata version published in 2018). *The IUCN Red List of Threatened Species 2016*.
- Ankel-Simons, F. (2007). *Primate Anatomy: An Introduction*. San Diego, CA: Academic Press.
- Antonelli, A., Hettling, H., Condamine, F. L., Vos, K., Nilsson, R. H., Sanderson, M. J., et al. (2016). Towards a Self-Updating Platform for Estimating Rates of Speciation and Migration, Ages, and Relationships of Taxa. *Systematic Biology*, 66(2), 152–166. <http://sysbio.oxfordjournals.org/>
- Appleby, B. M., Yamaguchi, N., Johnson, P. J., & Macdonald, D. W. (1999). Sex specific territorial responses in Tawny Owls *Strix aluco*. *Ibis*, 141, 91–99.
- Arbib, M. A., Liebal, K., & Pika, S. (2008). Primate Vocalization, Gesture, and the Evolution of Human Language. *Current Anthropology*, 49(6), 1053–1076. <https://doi.org/10.1086/593015>
- Arcadi, A. C. (1996). Phrase Structure of Wild Chimpanzee Pant Hoots: Patterns of Production and Interpopulation Variability. *American Journal of Primatology*, 39, 15–178. [https://doi.org/10.1002/\(SICI\)1098-2345\(1996\)39:3<159::AID-AJP2>3.0.CO;2-Y](https://doi.org/10.1002/(SICI)1098-2345(1996)39:3<159::AID-AJP2>3.0.CO;2-Y)
- Arcadi, A. C. (2000). Vocal responsiveness in male wild chimpanzees: Implications for the evolution of language. *Journal of Human Evolution*, 39(2), 205–223. <https://doi.org/10.1006/jhev.2000.0415>
- Arlet, M., Jubin, R., Masataka, N., & Lemasson, A. (2015). Grooming-at-a-distance by exchanging calls in non-human primates. *Biology Letters*, 11(10), 20150711–20150714. <https://doi.org/10.1098/rsbl.2015.0711>
- Arnold, K., & Whiten, A. (2003). Grooming interactions among the chimpanzees of the Budongo forest, Uganda: Tests of five explanatory models. *Behaviour*, 140, 519–552.
- Arnold, Kate, & Zuberbühler, K. (2006). Language evolution: Semantic combinations in primate calls. *Nature*, 441, 303. <https://doi.org/10.1038/441303a>
- Aroora, N., Nater, A., Van Schaik, C. P., Willems, E. P., Van Noordwijk, M. A., Goossens, B., et al. (2010). Effects of Pleistocene glaciations and rivers on the population structure of Bornean orangutans (*Pongo pygmaeus*). *Proceedings of the National Academy of Sciences of the United States of America*, 107(50), 21376–21381. <https://doi.org/10.1073/pnas.1010169107>
- Askew, J. A., & Morrogh-Bernard, H. C. (2017). Acoustic characteristics of long calls produced by male orang-utans (*Pongo pygmaeus wurmbii*): Advertising individual identity, context, and travel direction. *Folia Primatologica*, 87(5), 305–319. <https://doi.org/10.1159/000452304>
- Aureli, F., Schaffner, C. M., Boesch, C., Bearder, S. K., Call, J., Chapman, C. A., et al. (2008). Fission-fusion dynamics new research frameworks. *Current Anthropology*, 49(4), 627–654. <https://doi.org/10.1086/586708>
- Bailey, W. J. (2003). Insect duets: Underlying mechanisms and their evolution. *Physiological Entomology*, 28, 157–174. <https://doi.org/10.1046/j.1365-3032.2003.00337.x>
- Bailey, W. J., Hayasaka, K., Skinner, C. G., Kehoe, S., Sleu, L. C., Switom, J. L., & Goodman, M.

- (1992). Reexamination of the African Hominoid Trichotomy with Additional Sequences from the Primate P-Globin Gene Cluster. *Molecular Phylogenetics and Evolution*, 1(2), 97–135.
- Baker, T. M., Wilson, D. R., & Mennill, D. J. (2012). Vocal signals predict attack during aggressive interactions in black-capped chickadees. *Animal Behaviour*, 84, 965–974. <https://doi.org/10.1016/j.anbehav.2012.07.022>
- Baptista, L. F., & Gaunt, S. L. L. (1997). Social interaction and vocal development in birds. In C. T. Snowdon & M. Hausberger (Eds.), *Social influences on vocal development* (pp. 23–40). New York: Cambridge University Press.
- Beebe, B., Jaffe, J., Feldstein, S., Mays, K., & Alson, D. (1985). Interpersonal timing: the application of an adult dialogue model to mother-infant vocal and kinesic interactions. In T. Field and N. Fox (Ed.), *Social Perception in Infants* (pp. 217–247). Norwood, NJ: Ablex Publishing Corporation.
- Behr, O., Knörnschild, M., & Von Helversen, O. (2009). Territorial counter-singing in male sac-winged bats *saccopteryx bilineata*: Low-frequency songs trigger a stronger response. *Behavioral Ecology and Sociobiology*, 63(3), 433–442. <https://doi.org/10.1007/s00265-008-0677-2>
- Bergman, T. J. (2013). Speech-like vocalized lip-smacking in geladas. *Current Biology*, 23(7), R268–R269. <https://doi.org/10.1016/j.cub.2013.02.036>
- Bergman, T. J., & Beehner, J. C. (2015). Measuring social complexity. *Animal Behaviour*, 103, 203–209. <https://doi.org/10.1016/j.anbehav.2015.02.018>
- Bergman, Thore J, Beehner, J. C., Cheney, D. L., & Seyfarth, R. M. (2003). Hierarchical Classification by Rank and Kinship in Baboons Downloaded from. *Science*, 302(5648), 1234–1236. [www.sciencemag.org/cgi/content/full/302/5648/1231/](http://www.sciencemag.org/cgi/content/full/302/5648/1231/)
- Bermejo, M. (2004). Home-range use and intergroup encounters in western gorillas (*Gorilla g. gorilla*) at Lossi Forest, north Congo. *American Journal of Primatology*, 64(2), 223–232. <https://doi.org/10.1002/ajp.20073>
- Bermejo, M., Illera, G., & Sabater, J. (1994). Animals and Mushrooms Consumed by Bonobos (*Pan paniscus*): New Records from Lilungu (Ikela), Zaire. *International Journal of Primatology*, 15(6), 879–898.
- Bermejo, M., & Omedes, A. (1999). Preliminary Vocal Repertoire and Vocal Communication of Wild Bonobos (*Pan paniscus*) at Lilungu (Democratic Republic of Congo). *Folia Primatologica*, 70, 328–357.
- Bernardis, P., Bello, A., Pettenati, P., Stefanini, S., & Gentilucci, M. (2008). Manual actions affect vocalizations of infants. *Experimental Brain Research*, 184(4), 599–603. <https://doi.org/10.1007/s00221-007-1256-x>
- Bertin, A., Hausberger, M., Henry, L., & Richard-Yris, M. A. (2007). Adult and peer influences on starling song development. *Developmental Psychobiology*, 49(4), 362–374. <https://doi.org/10.1002/dev.20223>
- Biben, M., Symmes, D., & Masataka, N. (1986). Temporal and Structural Analysis of Affiliative Vocal Exchanges in Squirrel Monkeys (*Saimiri sciureus*). *Behaviour*, 98(1), 259–273.
- Bickerton, D. (1995). *Language and human behavior*. Seattle: University of Washington.
- Blumstein, D. T., & Armitage, K. B. (1997). Does sociality drive the evolution of communicative complexity? A comparative test with ground-dwelling sciurid alarm calls. *The American Naturalist*, 150, 179–200.
- Boesch, C. (1994). Cooperative Hunting in Wild Chimpanzees. *Animal Behaviour*, 48, 653–667.

- Boesch, C. (2002). Cooperative hunting roles among Tai chimpanzees. *Human Nature*, 13(1), 27–46.
- Boesch, C., & Boesch-Achermann, H. (2000). *The chimpanzees of the Tai forest: behavioural ecology and evolution*. Oxford University Press.
- Boesch, C., & Boesch, H. (1989). Hunting behavior of wild chimpanzees in the Tai National Park. *American Journal of Physical Anthropology*, 78, 547–573. <https://doi.org/10.1002/ajpa.1330780410>
- Boesch, C., Hohmann, G., & Marchant, L. F. (2002). *Behavioural diversity in chimpanzees and bonobos*. Cambridge, UK: Cambridge University Press.
- Bosch, J., & Márquez, R. (2000). Acoustical Interference in the Advertisement Calls of the Midwife Toads (*Alytes obstetricans* and *Alytes cisternasii*). *Behaviour*, 137(2), 249–263. <https://www.jstor.org/stable/4535702>
- Bouchet, H., Blois-Heulin, C., & Lemasson, A. (2013). Social complexity parallels vocal complexity: A comparison of three non-human primate species. *Frontiers in Psychology*, 4(390), 1–15. <https://doi.org/10.3389/fpsyg.2013.00390>
- Bouchet, H., Coye, C., & Lemasson, A. (2016). Le Langage est-il le Propre de l'Homme ? Apports des Etudes sur les Primates Non Humains. *Tétralogiques*, 21, 87–133.
- Bouchet, H., Koda, H., & Lemasson, A. (2017). Age-dependent change in attention paid to vocal exchange rules in Japanese macaques. *Animal Behaviour*, 129, 81–92. <https://doi.org/10.1016/j.anbehav.2017.05.012>
- Boughman, J. W., & Moss, C. F. (2003). Social sounds: vocal learning and development of mammal and bird calls. In A. . P. A. N. . F. R. Megela-Simmons (Ed.), *Acoustic communication* (pp. 138–224). New-York: Springer.
- Bradbury, J. W., & Vehrencamp, S. L. (1998). *Principles of animal communication*. Sunderland, MA: Sinauer Associates.
- Bradley, B. J., Doran-Sheehy, D. M., Lukas, D., Boesch, C., & Vigilant, L. (2004). Dispersed male networks in western gorillas. *Current Biology*, 14(6), 510–513. <https://doi.org/10.1016/j>
- Bradley, B. J., & Vigilant, L. (2002). The evolutionary genetics and molecular ecology of chimpanzees and bonobos. In *Behavioural diversity in chimpanzees and bonobos* (p. 259).
- Breuer, T., Robbins, A. M., Olejniczak, C., Parnell, R. J., Stokes, E. J., & Robbins, M. M. (2010). Variance in the male reproductive success of western gorillas: Acquiring females is just the beginning. *Behavioral Ecology and Sociobiology*, 64(4), 1–14. <https://doi.org/10.1007/s00265-009-0867-6>
- Breuer, T., Robbins, M. M., & Boesch, C. (2007). Using photogrammetry and color scoring to assess sexual dimorphism in wild western gorillas (*Gorilla gorilla*). *American Journal of Physical Anthropology*, 134(3), 369–382. <https://doi.org/10.1002/ajpa.20678>
- Briseño-Jaramillo, M., Ramos-Fernández, G., Palacios-Romo, T. M., Sosa-López, J. R., & Lemasson, A. (2018). Age and social affinity effects on contact call interactions in free-ranging spider monkeys. *Behavioral Ecology and Sociobiology*, 72, 192. <https://doi.org/10.1007/s00265-018-2615-2>
- Broca, P. (1865). Sur le siège de la faculté du langage articulé. *Bulletins de la Société Anthropologique de Paris*, 6, 377–393.
- Brockelman, W. Y., & Schilling, D. (1984). Inheritance of stereotyped gibbon calls'. *Nature*, 312, 634–636.
- Brown, E. D., & Farabaugh, S. M. (1991). Song Sharing in a Group-Living Songbird, the Australian



- Magpie, *Gymnorhina tibicen*. Part III. Sex Specificity and Individual Specificity of Vocal Parts in Communal Chorus and Duet Songs. *Source: Behaviour*, 118(3), 244–274.
- Brown, T. J., & Handford, P. (2003). Why birds sing at dawn: the role of consistent song transmission. *Ibis*, 145(1), 120–129.
- 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)
- Brumm, H., Voss, K., Köllmer, I., & Todt, D. (2003). Acoustic communication in noise: regulation of call characteristics in a New World monkey. *Journal of Experimental Biology*, 207, 443–448. <https://doi.org/10.1242/jeb.00768>
- Brunet, M., Guy, F., Pilbeam, D., Taisso Mackaye, H., Likius, A., Ahounta, D., et al. (2002). A new hominid from the Upper Miocene of Chad, Central Africa. *Nature*, 418(6894), 145–151. [www.nature.com/nature](http://www.nature.com/nature)
- Burkart, J. M., Guerreiro Martins, E., Miss, F., & Zürcher, Y. (2018). From sharing food to sharing information. *Interaction Studies*, 19(1–2), 136–150. <https://doi.org/10.1075/is.17026.bur>
- Burkart, J. M., Hrdy, S. B., & van Schaik, C. P. (2009). Cooperative breeding and human cognitive evolution. *Evolutionary Anthropology*, 18(5), 175–186. <https://doi.org/10.1002/evan.20222>
- Burkart, J. M., & van Schaik, C. P. (2010). Cognitive consequences of cooperative breeding in primates? *Animal Cognition*, 13(1), 1–19. <https://doi.org/10.1007/s10071-009-0263-7>
- Burt, J. M., & Vehrencamp, S. L. (2005). Dawn chorus as an interactive communication network. In P. K. McGregor (Ed.), *Animal communication networks* (pp. 320–344). Cambridge: Cambridge University Press.
- Busse, C. D. (1977). Chimpanzee Predation as a Possible Factor in the Evolution of Red Colobus Monkey Social Organization. *Evolution*, 31(4), 907–911.
- Byrne, R. W., Cartmill, E., Genty, E., Graham, K. E., Hobaiter, C., & Tanner, J. (2017). Great ape gestures: intentional communication with a rich set of innate signals. *Animal Cognition*, 20(4), 755–769. <https://doi.org/10.1007/s10071-017-1096-4>
- Caillaud, D., Levréro, F., Gatti, S., Ménard, N., & Raymond, M. (2008). Influence of male morphology on male mating status and behavior during interunit encounters in western lowland gorillas. *American Journal of Physical Anthropology*, 135(4), 379–388. <https://doi.org/10.1002/ajpa.20754>
- Calame-Griaule, G. (1965). *Ethnologie et langage. la parole chez les dogon*. Paris.
- Camacho-Schlenker, S., Courvoisier, H., & Aubin, T. (2011). Song sharing and singing strategies in the winter wren troglodytes troglodytes. *Behavioural Processes*, 87(3), 260–267. <https://doi.org/10.1016/j.beproc.2011.05.003>
- Candiotti, A., Zuberbuhler, K., & Lemasson, A. (2012). Convergence and divergence in Diana monkey vocalizations. *Biology Letters*, 8(3), 382–385. <https://doi.org/10.1098/rsbl.2011.1182>
- Cantalupo, C., & Hopkins, W. D. (2001). Asymmetric Broca's area in great apes. *Nature*, 414, 505.
- Caro, T. M., & Hauser, M. D. (1992). Is There Teaching in Nonhuman Animals? *The Quarterly Review of Biology*, 67(2), 151–174. <https://www.jstor.org/stable/2831436>
- Carter, G. G., Skowronski, M. D., Faure, P. A., & Fenton, B. (2008). Antiphonal calling allows individual discrimination in white-winged vampire bats. *Animal Behaviour*, 76, 1343–1355. <https://doi.org/10.1016/j.anbehav.2008.04.023>
- Casillas, M., Bobb, S. C., & Clark, E. V. (2016). Turn-taking, timing, and planning in early language

- acquisition. *Journal of Child Language*, 43(6), 1310–1337.  
<https://doi.org/10.1017/S0305000915000689>
- Castellano, D., & Munch, K. (2020). Populations Genomics in the Great apes. In Julien Y. Dutheil (Ed.), *Statistical Population Genomics* (Humana Press., pp. 453–468). Plön, Germany: Springer Protocols. <http://www.springer.com/series/7651>
- Chandrasekaran, C., Trubanova, A., Stillitano, S., Caplier, A., & Ghazanfar, A. A. (2009). The natural statistics of audiovisual speech. *PLoS Computational Biology*, 5(7), e1000436.  
<https://doi.org/10.1371/journal.pcbi.1000436>
- Chen, F.-C., & Li, W.-H. (2001). Genomic Divergences between Humans and Other Hominoids and the Effective Population Size of the Common Ancestor of Humans and Chimpanzees. *The American Journal of Human Genetics*, 68, 444–456.
- Chomsky, N. (1981). Knowledge of Language: Its Elements and Origins. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 295, 223–234.  
<https://doi.org/10.1098/rstb.1981.0135>
- Chow, C. P., Mitchell, J. F., & Miller, C. T. (2015). Vocal turn-taking in a non-human primate is learned during ontogeny. *Proceedings of the Royal Society B: Biological Sciences*, 282, 20150069. <https://doi.org/10.1098/rspb.2015.0069>
- Christie, P. J., Mennill, D. J., & Ratcliffe, L. M. (2004). Chickadee Song Structure Is Individually Distinctive over Long Broadcast Distances. *Behaviour*, 141(1), 101–124.
- Clark, A. P. (1993). Rank differences in the production of vocalizations by wild chimpanzees as a function of social context. *American Journal of Primatology*, 31, 159–179.
- Clark, A. P., & Wrangham, R. W. (1994). Chimpanzee Arrival Pant-Hoots: Do Signify Food or Status? *International Journal of Primatology*, 15(2), 185–205.
- Clark, C. W., & Clark, J. M. (1980). Sound Playback Experiments with Southern Right Whales (*Eubalaena australis*). *Science*, 207(8), 663–665.
- Clarke, E., Reichard, U. H., & Zuberbühler, K. (2006). The syntax and meaning of wild gibbon songs. *PLoS ONE*, 1(1), e73. <https://doi.org/10.1371/journal.pone.0000073>
- Clay, Z., Archbold, J., & Zuberbühler, K. (2015). Functional flexibility in wild bonobo vocal behaviour. *PeerJ*, 3, e1124. <https://doi.org/10.7717/peerj.1124>
- Clay, Z., Pika, S., Gruber, T., & Zuberbühler, K. (2011). Female bonobos use copulation calls as social signals. *Biology Letters*, 7(4), 513–516. <https://doi.org/10.1098/rsbl.2010.1227>
- Clay, Z., & Zuberbühler, K. (2009). Food-associated calling sequences in bonobos. *Animal Behaviour*, 77(6), 1387–1396. [www.Praat.org](http://www.Praat.org)
- Clay, Z., & Zuberbühler, K. (2011). Bonobos extract meaning from call sequences. *PLoS ONE*, 6(4), e18786. <https://doi.org/10.1371/journal.pone.0018786>
- Clay, Z., & Zuberbühler, K. (2012). Communication during sex among female bonobos: Effects of dominance, solicitation and audience. *Scientific Reports*, 2, 291.  
<https://doi.org/10.1038/srep00291>
- Cooksey, K., Sanz, C., Ebombi, T. F., Massamba, J. M., Teberd, P., Magera, E., et al. (2020). Socioecological Factors Influencing Intergroup Encounters in Western Lowland Gorillas (*Gorilla gorilla gorilla*). *International Journal of Primatology*, 41(2), 181–202.  
<https://doi.org/10.1007/s10764-020-00147-6>
- Corballis, M. C. (2002). From Hand to Mouth: The Gestural Origins of Language. In S. Christiansen, M. & Kirby (Ed.), *Language Evolution* (pp. 201–218). New York: Oxford University Press.

- Corballis, M. C. (2009). The evolution of language. *Annals of the New York Academy of Sciences*, 1156, 19–43. <https://doi.org/10.1111/j.1749-6632.2009.04423.x>
- Corballis, M. C. (2010). The gestural origins of language. *Wiley Interdisciplinary Reviews: Cognitive Science*, 1(1), 2–7. <https://doi.org/10.1002/wcs.2>
- Coudé, G., Ferrari, P. F., Rodà, F., Maranesi, M., Borelli, E., Veroni, V., et al. (2011). Neurons controlling voluntary vocalization in the macaque ventral premotor cortex. *PLoS ONE*, 6(11), e2682. <https://doi.org/10.1371/journal.pone.0026822>
- Coye, C., Townsend, S., & Lemasson, A. (2017). From animal communication to linguistics and back: insight from compositional abilities in monkeys and birds. In S. By Boë LJ, Fagot J, Perrier P & JL (Eds.), *Origins of human language: continuities and splits with nonhuman primates*. Peter Lang Publisher.
- Crockford, C., & Boesch, C. (2005). Call combinations in wild chimpanzees. *Behaviour*, 142(4), 397–421. <https://doi.org/10.1163/1568539054012047>
- Crockford, C., Herbinger, I., Vigilant, L., & Boesch, C. (2004). Wild Chimpanzees Produce Group-Specific Calls: a Case for Vocal Learning? *Ethology*, 110, 221–243.
- Crockford, C., Wittig, R. M., Mundry, R., & Zuberbühler, K. (2012). Wild chimpanzees inform ignorant group members of danger. *Current Biology*, 22(2), 142–146. <https://doi.org/10.1016/j.cub.2011.11.053>
- Crockford, C., Wittig, R. M., & Zuberbühler, K. (2015). An intentional vocalization draws others' attention: A playback experiment with wild chimpanzees. *Animal Cognition*, 18(3), 581–591. <https://doi.org/10.1007/s10071-014-0827-z>
- Da Cunha, R. G., & Byrne, R. W. (2006). Roars of Black Howler Monkeys (*Alouatta caraya*): Evidence for a Function in Inter-Group Spacing. *Behaviour*, 143(10), 1169–1199.
- Dabelsteen, T., McGregor, P. K., Holland, J., Tobias, J. A., & Pedersen, S. B. (1997). The signal function of overlapping singing in male robins. *Animal Behaviour*, 53(2), 249–256. <https://doi.org/10.1006/anbe.1996.0369>
- Das, R., Hergenrother, S. D., Soto-Calderón, I. D., Dew, J. L., Anthony, N. M., & Jensen-Seaman, M. I. (2014). Complete mitochondrial genome sequence of the eastern gorilla (*Gorilla beringei*) and implications for african ape biogeography. *Journal of Heredity*, 105(6), 752–761. <https://doi.org/10.1093/jhered/esu056>
- De Marco, A., Rebut, N., Massiot, E., Sanna, A., Sterck, E. H. M., Langermans, J. A. M., et al. (2019). Differential patterns of vocal similarity in tolerant and intolerant macaques. *Behaviour*, 1(aop), 1–25. <https://doi.org/10.1163/1568539X-00003562>
- De Waal, F. B. M. (1982). *Chimpanzee Politics*. (Jonathan Cape, Ed.). London.
- De Waal, F. B. M. (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>
- De Waal, F. B. M. (1995a). Primates: sex and societies. Collective matriarchat in Bonobo. *Sci Am*, 272(3), 82–88.
- De Waal, F. B. M. (1995b). Bonobo Sex and Society. *Scientific American*, 272(3), 82–88.
- De Waal, F. B. M. (1996). Macaque Social Culture: Development and Perpetuation of Affiliative Networks. *Journal of Comparative Psychology*, 110(2), 147–154.
- De Waal, F. B. M. (1997). *Bonobo: The Forgotten Ape*. Berkley, Los Angeles and London: University of California Press.

- Delgado, R. A., & van Schaik, C. P. (2000). The behavioral ecology and conservation of the orangutan (*Pongo pygmaeus*): A tale of two islands. *Evolutionary Anthropology: Issues, News, and Reviews*, 9(5), 1520–6505. [https://doi.org/10.1002/1520-6505\(2000\)9:5<201::AID-EVAN2>3.3.CO;2-P](https://doi.org/10.1002/1520-6505(2000)9:5<201::AID-EVAN2>3.3.CO;2-P)
- Dennett, D. C. (1983). Intentional systems in cognitive ethology: The “Panglossian paradigm” defended. *The Behavioral and Brain Sciences*, 6, 343–390.
- Digweed, S. M., Fedigan, L. M., & Rendall, D. (2007). Who cares who calls? Selective responses to the lost calls of socially dominant group members in the white-faced capuchin (*Cebus capucinus*). *American Journal of Primatology*, 69, 829–835. <https://doi.org/10.1002/ajp.20398>
- Dixson, A. (1998). *Primates sexuality*. (A. Bolin & P. Whelehan, Eds.). John Wiley & Sons.
- Dolata, J. K., Davis, B. L., & MacNeilage, P. F. (2008). Characteristics of the rhythmic organization of vocal babbling: Implications for an amodal linguistic rhythm. *Infant Behavior and Development*, 31(3), 422–431. <https://doi.org/10.1016/j.infbeh.2007.12.014>
- Dominguez, S., Devouche, E., Apter, G., & Gratier, M. (2016). The Roots of Turn-Taking in the Neonatal Period. *Infant and Child Development*, 25(3), 240–255. <https://doi.org/10.1002/icd.1976>
- Doran, D. M. (1996). Comparative positional behavior of the African apes. In *Great ape societies* (pp. 213–224).
- Duboscq, J., Neumann, C., Agil, M., Perwitasari-Farajallah, D., Thierry, B., & Engelhardt, A. (2017). Degrees of freedom in social bonds of crested macaque females. *Animal Behaviour*, 123, 411–426. <https://doi.org/10.1016/j.anbehav.2016.11.010>
- Duffy, K. G. (2006). *Social dynamics of male chimpanzees: adaptive significance of male bonds*. Los Angeles: University of California.
- Dunbar, R. I. M. (1993). Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Science*, 16, 681–735.
- Dunbar, R. I. M. (1996). *Grooming, gossip and the evolution of language*. London, United Kingdom: Faber & Faber.
- Dunbar, R. I. M. (1998). The Social Brain Hypothesis. *Evolutionary Anthropology*, 6(8), 178\_190.
- Dunbar, R. I. M. (2003). The origin and subsequent evolution of language. *Studies in the Evolution of Language*, 3, 219–234.
- Dunbar, R. I. M., Marriott, A., & Duncan, N. D. C. (1997). *Human Conversational behavior*.
- Duncan, S. (1972). Some signals and rules for taking speaking Turns in conversations. *Social Psychology*, 23(2), 283–292.
- Eaton, G. G., Johnson, D. F., Glick, B., & Worlein, J. M. (1985). Development in Japanese Macaques (*Macaca fuscata*)" Sexually Dimorphic Behavior During the First Year of Life. *Primates*, 26(3), 238–248.
- Elowson, M., Snowdon, C. T., & Sweet, C. . J. (1992). Ontogeny of trill and J-call vocalizations in the pygmy marmoset, *Cebuella pygmaea*. *Animal Behaviour*, 43, 703–715.
- Emery, N. (2000). The eye have it: The neuroethology, function and evolution of social gaze. *Neuroscience and Biobehavioral Reviews*, 24(6), 581–604. [https://doi.org/10.1016/S0149-7634\(00\)00025-7](https://doi.org/10.1016/S0149-7634(00)00025-7)
- Engesser, S., Ridley, A. R., & Townsend, S. W. (2016). Meaningful call combinations and compositional processing in the southern pied babbler. *Proceedings of the National Academy of*

- Sciences*, 113, 5976–5981. <https://doi.org/10.5061/dryad.ds0kb>
- Evans, C. S., & Evans, L. (1999). Chicken food calls are functionally referential. *Animal Behaviour*, 58, 307–319. <http://www.idealibrary.comon>
- Ey, E., & Fischer, J. (2009). The “acoustic adaptation hypothesis”—a review of the evidence from birds, anurans and mammals. *Bioacoustics*, 19(1–2), 21–48. <https://doi.org/10.1080/09524622.2009.9753613>
- Fabbri-Destro, M., & Rizzolatti, G. (2008). Mirror neurons and mirror systems in monkeys and humans. *American Physiological Society, Physiology*, 23(3), 15–24.
- Farabaugh, S. M. (1982). The ecological and social significance of duetting. In E. H. Kroodsma, D.E. & Miller (Ed.), *Acoustic communication in birds* (pp. 85–182). New York: Academic Press.
- Farabaugh, S. M., Brown, E. D., & Veltman, C. J. (1988). Song Sharing in a Group-Living Songbird, the Australian Magpie. Part II. Vocal Sharing between Territorial Neighbors, within and between Geographic Regions, and between Sexes. *Behaviour*, 104(1), 105–125.
- Faragó, T., Townsend, S., & Range, F. (2014). The information content of wolf (and dog) social communication. In *Biocommunication of Animals* (pp. 41–62). Springer Netherlands. [https://doi.org/10.1007/978-94-007-7414-8\\_4](https://doi.org/10.1007/978-94-007-7414-8_4)
- Fedurek, P., Donnellan, E., & Slocombe, K. E. (2014). Social and ecological correlates of long-distance pant hoot calls in male chimpanzees. *Behavioral Ecology and Sociobiology*, 68(8), 1345–1355. <https://doi.org/10.1007/s00265-014-1745-4>
- Fedurek, P., Machanda, Z. P., Schel, A. M., & Slocombe, K. E. (2013a). Pant hoot chorusing and social bonds in male chimpanzees. *Animal Behaviour*, 86(1), 189–196. <https://doi.org/10.1016/j.anbehav.2013.05.010>
- Fedurek, P., Schel, A. M., & Slocombe, K. E. (2013b). The acoustic structure of chimpanzee pant-hooting facilitates chorusing. *Behavioral Ecology and Sociobiology*, 67(11), 1781–1789. <https://doi.org/10.1007/s00265-013-1585-7>
- Fedurek, P., Slocombe, K. E., & Zuberbühler, K. (2015). Chimpanzees communicate to two different audiences during aggressive interactions. *Animal Behaviour*, 110, 21–28. <https://doi.org/10.1016/j.anbehav.2015.09.010>
- Fedurek, P., Zuberbühler, K., & Dahl, C. D. (2016). Sequential information in a great ape utterance. *Scientific Reports*, 6, 38226. <https://doi.org/10.1038/srep38226>
- Fischer, J., Farnworth, M. S., Sennhenn-Reulen, H., & Hammerschmidt, K. (2017a). Quantifying social complexity. *Animal Behaviour*, 130, 57–66. <https://doi.org/10.1016/j.anbehav.2017.06.003>
- Fischer, J., Kitchen, D. M., Seyfarth, R. M., & Cheney, D. L. (2004). Baboon loud calls advertise male quality: Acoustic features and their relation to rank, age, and exhaustion. *Behavioral Ecology and Sociobiology*, 56, 140–148. <https://doi.org/10.1007/s00265-003-0739-4>
- Fischer, J., Wadewitz, P., & Hammerschmidt, K. (2017b). Structural variability and communicative complexity in acoustic communication. *Animal Behaviour*, 134, 229–237. <https://doi.org/10.1016/j.anbehav.2016.06.012>
- Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in Cognitive Sciences*, 4(7), 258–267.
- Fitch, W. T. (2010). *The Evolution of Language*. Cambridge: Cambridge University Press.
- Foote, A. D., Griffin, R. M., Howitt, D., Larsson, L., Miller, P. J. O., & Hoelzel, A. R. (2006). Killer whales are capable of vocal learning. *Biology Letters*, 2(4), 509–512. <https://doi.org/10.1098/rsbl.2006.0525>

- Foote, J. R., Fitzsimmons, L. P., Mennill, D. J., & Ratcliffe, L. M. (2011). Male black-capped chickadees begin dawn chorusing earlier in response to simulated territorial insertions. *Animal Behaviour*, *81*(4), 871–877. <https://doi.org/10.1016/j.anbehav.2011.01.028>
- Forcina, G., Vallet, D., Le Gouar, P. J., Bernardo-Madrid, R., Illera, G., Molina-Vacas, G., et al. (2019). From groups to communities in western lowland gorillas. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1896), n/a. <https://doi.org/10.1098/rspb.2018.2019>
- Fouquet, A., Tilly, T., Pašukonis, A., Courtois, E. A., Gaucher, P., Ulloa, J., & Sueur, J. (2020). Simulated chorus attracts conspecific and heterospecific Amazonian explosive breeding frogs. *Biotropica*, *00*, 1–11. <https://doi.org/10.1111/btp.12845>
- Freeberg, T. M. (2006). Social Complexity Can Drive Vocal Complexity Group Size Influences Vocal Information in Carolina Chickadees. *Psychological Science*, *17*, 557–561.
- Freeberg, T. M., Dunbar, R. I. M., & Ord, T. J. (2012). Social complexity as a proximate and ultimate factor in communicative complexity. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*, 1785–1801. <https://doi.org/10.1098/rstb.2011.0213>
- Freeberg, T. M., & Lucas, J. R. (2012). Information theoretical approaches to chick-a-dee calls of carolina chickadees (*Poecile carolinensis*). *Journal of Comparative Psychology*, *126*(1), 68–81. <https://doi.org/10.1037/a0024906>
- Fröhlich, M., Kuchenbuch, P., Müller, G., Fruth, B., Furuichi, T., Wittig, R. M., & Pika, S. (2016). Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. *Scientific Reports*, *6*(25887), 1–14. <https://doi.org/10.1038/srep25887>
- Fröhlich, M., Sievers, C., Townsend, S. W., Gruber, T., & van Schaik, C. P. (2019). Multimodal communication and language origins: integrating gestures and vocalizations. *Biological Reviews*, *94*(5), 1809–1829. <https://doi.org/10.1111/brv.12535>
- Fröhlich, M., & van Schaik, C. P. (2018). The function of primate multimodal communication. *Animal Cognition*, *21*(5), 619–629. <https://doi.org/10.1007/s10071-018-1197-8>
- Fruth, B., Hickey, J. R., André, C., Furuichi, T., Hart, J., Hart, T., et al. (2016). *Pan paniscus* (errata version published in 2016). *The IUCN Red List of Threatened Species 2016*.
- Fruth, B., & Hohmann, G. (2002). How bonobos handle hunts and harvests: why share food? In G. H. & L. M. C. Boesch (Ed.), *Behavioural Diversity in Chimpanzees and Bonobos* (pp. 231–243). Cambridge, UK: Cambridge University Press.
- Fruth, B., Williamson, E. A., & Richardson, M. C. (2013). Bonobo *Pan paniscus*. In A. B. R. and D. E. W. R.A. Mittermeier (Ed.), *Handbook of the Mammals of the World Volume 3: Primates* (Lynx Edicions., pp. 853–854). Barcelona, Spain.
- Funkhouser, J. A., Mayhew, J. A., & Mulcahy, J. B. (2018). Social network and dominance hierarchy analyses at Chimpanzee Sanctuary Northwest. *PlosOne*, *13*(2), e0191898. <https://doi.org/10.1371/journal.pone.0191898>
- Furrer, R. D., & Manser, M. B. (2009). The evolution of urgency-based and functionally referential alarm calls in ground-dwelling species. *American Naturalist*, *173*(3), 400–410. <https://doi.org/10.1086/596541>
- Furuichi, T. (1989). Social Interactions and the Life History of Female *Pan paniscus* in Wamba, Zaire. *International Journal of Primatology*, *10*(3), 173–197.
- Furuichi, T. (2011). Female contributions to the peaceful nature of bonobo society. *Evolutionary Anthropology*, *20*, 131–142. <https://doi.org/10.1002/evan.20308>
- Furuichi, T., Idani, G., Ihobe, H., Hashimoto, C., Tashiro, Y., Sakamaki, T., et al. (2012). Long-term studies on wild bonobos at Wamba, Luo Scientific Reserve, D. R. Congo: Towards the

- understanding of female life history in a male-philopatric species. In Heidelberg (Ed.), *Long-Term Field Studies of Primates* (Springer., Vol. 9783642225147, pp. 413–433). Berlin: Springer-Verlag Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-22514-7\\_18](https://doi.org/10.1007/978-3-642-22514-7_18)
- Furuichi, T., & Ihobe, H. (1994). Variation in male relationships in bonobos and chimpanzees. *Behaviour*, 130, 211–228.
- Galdikas, B. M. F. (1984). Adult female sociality among wild orangutans at Tanjung Puting Reserve. In M. Small (Ed.), *Female Primates: Studies by Women Primatologists* (pp. 217–235). New York: Liss.
- Galdikas, B. M. F. (1985a). Orangutan Sociality at Tanjung Puting. *American Journal of Primatology*, 9, 119.
- Galdikas, B. M. F. (1985b). Subadult Male Orangutan Sociality and Reproductive Behavior at Tanjung Puting. *American Journal of Primatology*, 8(2), 8237–99.
- Galdikas, B. M. F. (1985c). Adult male sociality and reproductive tactics among orangutans at Tanjung Puting. *Folia Primatologica*, 45(9), 9–24.
- Gardner, R. A., & Gardner, B. T. (1969). Teaching Sign Language to a Chimpanzee. *Science*, 165, 664–672.
- Gatti, S., Levréro, F., Ménard, N., & Gautierhion, A. (2004). Population and group structure of western Lowland Gorillas (*Gorilla gorilla gorilla*) at Lokoué, Republic of Congo. *American Journal of Primatology*, 63, 111–123. <https://doi.org/10.1002/ajp.20045>
- Gautier, J.-P. (2001). Communication acoustique et environnement chez les primates simiens. *Primatologie*, 4, 49–74.
- Geissmann, T. (1984). Separatum Inheritance of Song Parameters in the Gibbon Song, Analysed in 2 Hybrid Gibbons (*Hylobates pileatus* X *H. lar*). *Folia Primatologica*, 42, 216–235.
- Geissmann, T. (2002). Duet-splitting and the evolution of gibbon songs. *Biological Reviews of the Cambridge Philosophical Society*, 77, 57–76. <https://doi.org/10.1017/S1464793101005826>
- Geissmann, T., & Mutschler, T. (2006). Diurnal distribution of loud calls in sympatric wild indris (*indri indri*) and ruffed lemurs (*varecia variegata*): Implications for call functions. *Primates*, 47(4), 393–396. <https://doi.org/10.1007/s10329-006-0189-5>
- Gentilucci, M., & Volta, R. D. (2008). Spoken language and arm gestures are controlled by the same motor control system. *The Quarterly Journal of Experimental Psychology*, 61(6), 944–957.
- Genty, E., Breuer, T., Hobaiter, C., & Byrne, R. W. (2009). Gestural communication of the gorilla (*Gorilla gorilla*): Repertoire, intentionality and possible origins. *Animal Cognition*, 12, 527–546. <https://doi.org/10.1007/s10071-009-0213-4>
- Genty, E., Clay, Z., Hobaiter, C., & Zuberbühler, K. (2014). Multi-modal use of a socially directed call in bonobos. *PLoS ONE*, 9(1), e84738. <https://doi.org/10.1371/journal.pone.0084738>
- Genty, E., & Zuberbühler, K. (2014). Spatial reference in a bonobo gesture. *Current Biology*, 24(14), 1601–1605. <https://doi.org/10.1016/j.cub.2014.05.065>
- Ghazanfar, A. A., Chandrasekaran, C., & Morrill, R. J. (2010). Dynamic, rhythmic facial expressions and the superior temporal sulcus of macaque monkeys: Implications for the evolution of audiovisual speech. *European Journal of Neuroscience*, 31, 1807–1817. <https://doi.org/10.1111/j.1460-9568.2010.07209.x>
- Ghazanfar, A. A., & Logothetis, N. K. (2002). Facial expressions linked to monkey calls. *Nature*, 423, 937–938.

- Ghazanfar, A. A., & Rendall, D. (2008). Evolution of human vocal production. *Current Biology*, 18(11), R457–R460. <https://doi.org/10.1016/j.cub.2008.03.030>
- Ghazanfar, A. A., & Takahashi, D. Y. (2014). The evolution of speech: Vision, rhythm, cooperation. *Trends in Cognitive Sciences*, 18(10), 543–553. <https://doi.org/10.1016/j.tics.2014.06.004>
- Ghazanfar, A. A., Takahashi, D. Y., Mathur, N., & Fitch, W. T. (2012). Cineradiography of monkey lip-smacking reveals putative precursors of speech dynamics. *Current Biology*, 22, 1176–1182. <https://doi.org/10.1016/j.cub.2012.04.055>
- Gilby, I. C., & Wrangham, R. W. (2008). Association patterns among wild chimpanzees (*Pan troglodytes schweinfurthii*) reflect sex differences in cooperation. *Behavioral Ecology and Sociobiology*, 62(11), 1831–1842. <https://doi.org/10.1007/s00265-008-0612-6>
- Gillespie-Lynch, K., Greenfield, P., Lyn, H., & Savage-Rumbaugh, S. (2014). Gestural and symbolic development among apes and humans: Support for a multimodal theory of language evolution. *Frontiers in Psychology*, 5(OCT). <https://doi.org/10.3389/fpsyg.2014.01228>
- Ginsburg, G. P., & Kilbourne, B. K. (1988). Emergence of vocal alternation in mother-infant interchanges. *Journal of Child Language*, 15, 221–235. <https://doi.org/10.1017/S0305000900012344>
- Girard-Buttoz, C., Surbeck, M., Samuni, L., Boesch, C., Fruth, B., Crockford, C., et al. (2020). Variable use of polyadic grooming and its effect on access to social partners in wild chimpanzees and bonobos. *Animal Behaviour*, 168, 211–224. <https://doi.org/10.1016/j.anbehav.2020.08.021>
- Goldberg, J. A. (1990). Interrupting the discourse on interruptions - an analysis in terms of relationally neutral, power-oriented and rapport-oriented acts. *Journal of Pragmatics*, 14(6), 883–903.
- Goll, Y., Demartsev, V., Koren, L., & Geffen, E. (2017). Male hyraxes increase countersinging as strangers become ‘nasty neighbours.’ *Animal Behaviour*, 134, 9–14. <https://doi.org/10.1016/j.anbehav.2017.10.002>
- Gómez, J. M., Verdú, M., González-Megías, A., & Méndez, M. (2016). The phylogenetic roots of human lethal violence. *Nature*, 538(7624), 233–237. <https://doi.org/10.1038/nature19758>
- Gonder, M. K., Locatelli, S., Ghobrial, L., Mitchell, M. W., Kujawski, J. T., Lankester, F. J., et al. (2011). Evidence from Cameroon reveals differences in the genetic structure and histories of chimpanzee populations. *Proceedings of the National Academy of Sciences of the United States of America*, 108(12), 4766–4771. <https://doi.org/10.1073/pnas.1015422108>
- Goodall, J. (1986). *The Chimpanzees of Gombe: Patterns of Behavior*. Cambridge: Belknap Press of Harvard University Press.
- Gouzoules, H., & Gouzoules, S. (1995). Recruitment Screams of Pigtail Monkeys (*Macaca nemestrina*): Ontogenetic Perspectives. *Behaviour*, 132(5), 431–450.
- Gouzoules, H., & Gouzoules, S. (1989). Design features and developmental modification of pigtail macaque, *Macaca nemestrina*, agonistic screams. *Animal Behaviour*, 37, 383–401.
- Gouzoules, S., Gouzoules, H., & Marler, P. (1984). Rhesus Monkey (*Macaca mulata*) Screams: Representational Signaling in the Recruitment of Agonistic Aid. *Animal Behaviour*, 32, 182–193.
- Grafe, T. U. (1996). The function of call alternation in the African reed frog (*Hyperolius marmoratus*): precise call timing prevents auditory masking. *Behavioral Ecology and Sociobiology*, 38, 149–158.
- Graham, K. E., Furuichi, T., & Byrne, R. W. (2017). The gestural repertoire of the wild bonobo (*Pan paniscus*): a mutually understood communication system. *Animal Cognition*, 20(2), 171–177. <https://doi.org/10.1007/s10071-016-1035-9>



- Graham, K. E., Hobaiter, C., Ounsley, J., Furuichi, T., & Byrne, R. W. (2018). Bonobo and chimpanzee gestures overlap extensively in meaning. *PLoS Biology*, *16*(2).  
<https://doi.org/10.1371/journal.pbio.2004825>
- Graham, K. E., Wilke, C., Lahiff, N. J., & Slocombe, K. E. (2019). Scratching beneath the surface: Intentionality in great ape signal production. *Philosophical Transactions of the Royal Society B: Biological Sciences*. <https://doi.org/10.1098/rstb.2018.0403>
- Gratier, M., Devouche, E., Guellai, B., Infanti, R., Yilmaz, E., & Parlato-Oliveira, E. (2016). Early development of turn-taking in vocal interaction between mothers and infants. In J. Holler, K. H. Kendrick, M. Casillas, & S. C. Levinson (Eds.), *Turn-Taking in Human Communicative Interaction* (Vol. 6, pp. 236–245). Frontiers in Psychology.
- Graw, B., & Manser, M. B. (2007). The function of mobbing in cooperative meerkats. *Animal Behaviour*, *74*(3), 507–517. <https://doi.org/10.1016/j.anbehav.2006.11.021>
- Greenberg, S., Carvey, H., Hitchcock, L., & Chang, S. (2003). Temporal properties of spontaneous speech - A syllable-centric perspective. *Journal of Phonetics*, *31*(3–4), 465–485.  
<https://doi.org/10.1016/j.wocn.2003.09.005>
- Greenfield, M. D. (1994). Cooperation and Conflict in the Evolution of Signal Interactions. *Annual Review of Ecology, Evolution, and Systematics*, *25*, 97–126.
- Greenfield, M. D. (2014). Signal interactions and interference in insect choruses: singing and listening in the social environment. *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, *201*(1), 143–154. <https://doi.org/10.1007/s00359-014-0938-7>
- Greenfield, M. D., & Minckley, R. L. (1993). Acoustic Dueling in Tarbush Grasshoppers: Settlement of Territorial Contests via Alternation of Reliable Signals. *Ethology*, *95*, 309–326.
- Greenfield, M. D., & Rand, A. S. (2000). Frogs have rules: selective attention algorithms regulate chorusing in *Physalaemus pustulosus* (Leptodactylidae). *Ethology*, *106*, 331–347.
- Greenfield, M. D., & Roizen, I. (1993). Katydid synchronous chorusing is an evolutionarily stable outcome of female choice. *Nature*, *364*, 618–620.
- Greminger, M. P., Stölting, K. N., Nater, A., Goossens, B., Arora, N., Bruggmann, R., et al. (2014). Generation of SNP datasets for orangutan population genomics using improved reduced-representation sequencing and direct comparisons of SNP calling algorithms. *BMC Genomics*, *15*, 16.
- Groves, C. P. (2001). *Primate Taxonomy*. Washington: Smithsonian Institution Press.
- Gruber, T., & Zuberbühler, K. (2013). Vocal Recruitment for Joint Travel in Wild Chimpanzees. *PLoS ONE*, *8*(9), e76073. <https://doi.org/10.1371/journal.pone.0076073>
- Gustison, M. L., & Bergman, T. J. (2017). Divergent acoustic properties of gelada and baboon vocalizations and their implications for the evolution of human speech. *Journal of Language Evolution*, *2*(1), 20–36. <https://doi.org/10.1093/jole/lzx015>
- Gustison, M. L., le Roux, A., & Bergman, T. J. (2012). Derived vocalizations of geladas (*Theropithecus gelada*) and the evolution of vocal complexity in primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1597), 1847–1859.  
<https://doi.org/10.1098/rstb.2011.0218>
- Hage, S. R., & Nieder, A. (2013). Single neurons in monkey prefrontal cortex encode volitional initiation of vocalizations. *Nature Communications*, *4*, 2409.  
<https://doi.org/10.1038/ncomms3409>
- Hagemann, L., Boesch, C., Robbins, M. M., Arandjelovic, M., Deschner, T., Lewis, M., et al. (2018).

- Long-term group membership and dynamics in a wild western lowland gorilla population (Gorilla gorilla gorilla) inferred using non-invasive genetics. *American Journal of Primatology*, 80(8). <https://doi.org/10.1002/ajp.22898>
- Haimoff, E. H. (1986). Convergence in the duetting of monogamous Old World primates. *Journal of Human Evolution*, 15(1), 51–59. [https://doi.org/10.1016/S0047-2484\(86\)80065-3](https://doi.org/10.1016/S0047-2484(86)80065-3)
- Hall, M. L. (2000). The function of duetting in magpie-larks: Conflict, cooperation, or commitment? *Animal Behaviour*, 60(5), 667–677. <https://doi.org/10.1006/anbe.2000.1517>
- Hall, M. L. (2004). A review of hypotheses for the functions of avian duetting. *Behavioral Ecology and Sociobiology*, 55(5), 415–430. <https://doi.org/10.1007/s00265-003-0741-x>
- Hammerschmidt, K., & Fischer, J. (1998). The Vocal Repertoire of Barbary Macaques: A Quantitative Analysis of a Graded Signal System. *Ethology*, 104, 203–216.
- Hammerschmidt, K., & Fischer, J. (2008). Constraints in primate vocal production. In O. K. Griebel U (Ed.), *The Evolution of Communicative Creativity: From Fixed Signals to Contextual Flexibility* (pp. 93–120). MIT Press.
- Hammerschmidt, K., & Fischer, J. (2019). Baboon vocal repertoires and the evolution of primate vocal diversity. *Journal of Human Evolution*, 126, 1–13. <https://doi.org/10.1016/j.jhevol.2018.10.010>
- Harcourt, A. H., & Stewart, K. J. (2007). Gorilla society: What we know and don't know. *Evolutionary Anthropology*, 16(4), 147–158. <https://doi.org/10.1002/evan.20142>
- Harcourt, A., Stewart, K., & Hauser, M. (1993). Functions of wild gorilla 'close' calls. I. Repertoire, context, and interspecific comparison. *Behaviour*, 124(1), 89–122. <https://doi.org/10.1017/CBO9781107415324.004>
- Hardus, M. E., Lameira, A. R., Singleton, I., Morrogh-Bernard, H. C., Knott, C. D., Ancrenaz, M., et al. (2009). A description of the orangutan's vocal and sound repertoire, with a focus on geographic variation. In Oxford University Press (Ed.), *Orangutans: Geographic Variation in Behavioral Ecology and Conservation* (pp. 49–64). New-York. <https://doi.org/10.1093/acprof:oso/9780199213276.003.0004>
- Harrington, F. H., & David, L. (1979). *Wolf Howling and Its Role in Territory Maintenance*. *Mech Source: Behaviour* (Vol. 68).
- Hartbauer, M., Kratzer, S., Steiner, K., & Römer, H. (2005). Mechanisms for synchrony and alternation in song interactions of the bushcricket *Mecopoda elongata* (Tettigoniidae: Orthoptera). *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology*, 191(2), 175–188. <https://doi.org/10.1007/s00359-004-0586-4>
- Hartbauer, M., & Römer, H. (2016). Rhythm generation and rhythm perception in insects: The evolution of synchronous choruses. *Frontiers in Neuroscience*, 10. <https://doi.org/10.3389/fnins.2016.00223>
- Hashimoto, C., Tashiro, Y., Kimura, D., Enomoto, T., Ingmanson, E. J., Idani, G. 'I., & Furuichi, T. (1998). Habitat Use and Ranging of Wild Bonobos (*Pan paniscus*) at Wamba. *International Journal of Primatology*, 19(6), 1045–1060.
- Hausberger, M., Richard-Yris, M.-A., Henry, L., Lepage, L., & Schmidt, I. (1995). Song Sharing Reflects the Social Organization in a Captive Group of European Starlings (*Sturnus vulgaris*). *Journal of Comparative Psychology*, 109(3), 222–241.
- Hauser, M. D. (1989). Ontogenetic Changes in the Comprehension and Production of Vervet Monkey (*Cercopithecus aethiops*) Vocalizations. *Journal of Comparative Psychology*, 103, 2–149.
- Hauser, M. D. (1996). *The evolution of communication*. Cambridge, MA: MIT Press.

- Hayashi, M. (2013). Turn allocation and turn sharing. In *Handbook of Conversation Analysis* (pp. 167–190). Chichester: Wiley-Blackwell.
- Hedwig, D., Mundry, R., Robbins, M. M., & Boesch, C. (2015a). Contextual correlates of syntactic variation in mountain and western gorilla close-distance vocalizations: Indications for lexical or phonological syntax? *Animal Cognition*, *18*, 423–435. <https://doi.org/10.1007/s10071-014-0812-6>
- Hedwig, D., Mundry, R., Robbins, M. M., & Boesch, C. (2015b). Audience effects, but not environmental influences, explain variation in gorilla close distance vocalizations-A test of the acoustic adaptation hypothesis. *American Journal of Primatology*, *77*, 1239–1252. <https://doi.org/10.1002/ajp.22462>
- Hedwig, D., Robbins, M. M., Mundry, R., Hammerschmidt, K., & Boesch, C. (2014). Acoustic structure and variation in mountain and western gorilla close calls: A syntactic approach. *Behaviour*, *151*, 1091–1120. <https://doi.org/10.1163/1568539X-00003175>
- Henry, L., Craig, A. J. F. K., Lemasson, A., & Hausberger, M. (2015). Social coordination in animal vocal interactions. Is there any evidence of turn-taking? The starling as an animal model. *Frontiers in Psychology*, *6*, 1416. <https://doi.org/10.3389/fpsyg.2015.01924>
- Hewes, G. W. (1973). Primate communication and the gestural origin of language. *Current Anthropology*, *14*(1–2), 65–84.
- Hilbrink, E. E., Gattis, M., & Levinson, S. C. (2015). Early developmental changes in the timing of turntaking: a longitudinal study of mother-infant interaction. *Frontiers in psychology*, *6*, 1492. <https://doi.org/10.3389/978-2-88919-825-2>
- Hinde, R. A. (1976). Interactions, relationships and social structure. *Man*, *11*, 1–17. <https://about.jstor.org/terms>
- Hobaiter, C., & Byrne, R. W. (2011). The gestural repertoire of the wild chimpanzee. *Animal Cognition*, *14*(5), 745–767. <https://doi.org/10.1007/s10071-011-0409-2>
- Hobaiter, C., & Byrne, R. W. (2013). Flexibilité et intentionnalité dans la communication gestuelle chez les grands singes | Flexibility and intentionality in non-human great ape gestural communication. *Revue de primatologie*, (5). <https://doi.org/10.4000/primatologie.1713>
- Hobaiter, C., Byrne, R. W., & Zuberbühler, K. (2017). Wild chimpanzees' use of single and combined vocal and gestural signals. *Behavioral Ecology and Sociobiology*, *71*(6). <https://doi.org/10.1007/s00265-017-2325-1>
- Hobolth, A., Dutheil, J. Y., Hawks, J., Schierup, M. H., & Mailund, T. (2011). Incomplete lineage sorting patterns among human, chimpanzee, and orangutan suggest recent orangutan speciation and widespread selection. *Genome Research*, *21*(3), 349–356. <https://doi.org/10.1101/gr.114751.110>
- Hockett, C. F., & Hockett, C. D. (1960). The Origin of Speech. *Scientific American*, *203*(3), 88–97. <https://doi.org/10.2307/24940617>
- Hohmann, G. (2001). Association and Social Interactions Between Strangers and Residents in Bonobos (*Pan paniscus*). *Primates*, *42*(1), 91–99.
- Hohmann, G., & Fruth, B. (2000). Use and function of genital contacts among female bonobos. *Animal Behaviour*, *60*(1), 107–120. <https://doi.org/10.1006/anbe.2000.1451>
- Hohmann, G., & Fruth, B. (2008). New records on prey capture and meat eating by bonobos at Lui Kotale, Salonga National Park, Democratic Republic of Congo. *Folia Primatologica*, *79*(2), 103–110. <https://doi.org/10.1159/000110679>
- Hopkins, W. D., Tagliatalata, J. P., & Leavens, D. A. (2007). Chimpanzees differentially produce

- novel vocalizations to capture the attention of a human. *Animal Behaviour*, 73(2), 281–286. <https://doi.org/10.1016/j.anbehav.2006.08.004>
- Horr, D. A. (1975). The Borneo orang-utan: Population structure and dynamics in relationship to ecology and reproductive strategy. In L. Rosenblum (Ed.), *Primate Behavior. Developments in Field and Laboratory Research* (pp. 307–323). New York: Academic Press.
- Horwitz, B., Amunts, K., Bhattacharyya, R., Patkin, D., Jeffries, K., Zilles, K., & Braun, A. R. (2003). Activation of Broca's area during the production of spoken and signed language: A combined cytoarchitectonic mapping and PET analysis. *Neuropsychologia*, 41(14), 1868–1876. [https://doi.org/10.1016/S0028-3932\(03\)00125-8](https://doi.org/10.1016/S0028-3932(03)00125-8)
- Idani, G. (1991). Social relationships between immigrant and resident bonobo (*Pan paniscus*) females at Wamba. *Folia Primatologica*, 57, 83–95.
- Idani, G., Kuroda, S., Kano, T., & Asato, R. (1994). Flora and vegetation of Wamba forest, Central Zaire with reference to bonobo (*Pan paniscus*) foods. *Tropics*, 3, 309–332.
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., & Jasnow, M. D. (2001). Rhythms of dialogue in infancy: coordinated timing in development. *Monographs of the Society for Research in Child Development*, 66(2), 1–132. <http://www.ncbi.nlm.nih.gov/pubmed/11428150>
- Janik, V. M. (2000). Whistle matching in wild bottlenose dolphins (*Tursiops truncatus*). *Science*, 289(5483), 1355–1357. <https://doi.org/10.1126/science.289.5483.1355>
- Janik, V. M., & Slater, P. J. (1997). Vocal learning in mammals. *Advances in the Study of Behavior*, 26, 59–100.
- Jarvis, E. D. (2004). Learned Birdsong and the Neurobiology of Human Language. *Annals of the New York Academy of Science*, 1016, 749–777. <http://www.jarvislab.net/>
- Jordan, K., Weiss, D., Hauser, M., & McMurray, B. (2004). Antiphonal Responses to Loud Contact Calls Produced by *Saguinus oedipus*. *International Journal of Primatology*, 25(2), 465–475.
- Jürgens, U. (1995). Neuronal Control of Vocal Production in Non-Human and Human Primates. In J. U. Zimmermann E., Newman J.D. (Ed.), *Current Topics in Primate Vocal Communication* (pp. 199–206). Springer, Boston, MA. [https://doi.org/10.1007/978-1-4757-9930-9\\_10](https://doi.org/10.1007/978-1-4757-9930-9_10)
- Kacelnik, A., & Krebs, J. R. (1982). The Dawn Chorus in the Great Tit (*Parus major*): Proximate and Ultimate Causes. *Behaviour*, 83(4), 287–309.
- Kanamori, T., Kuze, N., Bernard, H., Malim, T. P., & Kohshima, S. (2010). Feeding ecology of Bornean orangutans (*Pongo pygmaeus morio*) in Danum Valley, Sabah, Malaysia: A 3-year record including two mast fruitings. *American Journal of Primatology*, 72(9), 820–840. <https://doi.org/10.1002/ajp.20848>
- Kano, T. (1982). The Social Group of Pygmy Chimpanzees (*Pan paniscus*) of Wamba. *Primates*, 23(2), 171–188.
- Kano, T. (1992). *The last ape: Pygmy chimpanzee behavior and ecology*. Stanford, California: Stanford University Press.
- Kaplan, H. S., Hooper, P. L., & Gurven, M. (2009). The evolutionary and ecological roots of human social organization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1533), 3289–3299. <https://doi.org/10.1098/rstb.2009.0115>
- Kappeler, P. M., & van Schaik, C. P. (2002). Evolution of Primate Social Systems. *International Journal of Primatology*, 23(4), 707–740.
- Katsu, N., Yamada, K., & Nakamichi, M. (2014). Development in the usage and comprehension of greeting calls in a free-ranging group of Japanese macaques (*Macaca fuscata*). *Ethology*,

- 120(10), 1024–1034. <https://doi.org/10.1111/eth.12275>
- Keenan, S., Mathevon, N., Stevens, J. M. G., Nicolè, F., Zuberbühler, K., Guéry, J.-P., & Levréro, F. (2020). The reliability of individual vocal signature varies across the bonobo's graded repertoire. *Animal Behaviour*, *169*, 9–21. <https://doi.org/10.1016/j.anbehav.2020.08.024>
- Keenan, S., Mathevon, N., Stevens, J. M., Guéry, J. P., Zuberbühler, K., & Levréro, F. (2016). Enduring voice recognition in bonobos. *Scientific Reports*, *6*. <https://doi.org/10.1038/srep22046>
- Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press.
- Kendon, A. (2017). Reflections on the “gesture-first” hypothesis of language origins. *Psychonomic Bulletin and Review*, *24*(1), 163–170. <https://doi.org/10.3758/s13423-016-1117-3>
- Kirschel, A. N. G., Blumstein, D. T., Cohen, R. E., Buermann, W., Smith, T. B., & Slabbekoorn, H. (2009). Birdsong tuned to the environment: Green hylia song varies with elevation, tree cover, and noise. *Behavioral Ecology*, *20*(5), 1089–1095. <https://doi.org/10.1093/beheco/arp101>
- Kitchen, D. M., Seyfarth, R. M., Fischer, J., & Cheney, D. L. (2003). Loud calls as indicators of dominance in male baboons (*Papio cynocephalus ursinus*). *Behavioral Ecology and Sociobiology*, *53*(6), 374–384. <https://doi.org/10.1007/s00265-003-0588-1>
- Knott, C., Beaudrot, L., Snaith, T., White, S., Tschauner, H., & Planansky, G. (2008). Female-female competition in Bornean orangutans. *International Journal of Primatology*, *29*, 975–997. <https://doi.org/10.1007/s10764-008-9278-1>
- Knott, C. D. (1999). Orangutan behavior and ecology. In A. Dolhinow, P., and Fuentes (Ed.), *The nonhuman primates* (pp. 50–57). Mountain View: Mayfield Publishing.
- Koda, H. (2004). Flexibility and context-sensitivity during the vocal exchange of coo calls in wild Japanese macaques (*Macaca fuscata yakui*). *Behaviour*, *141*(10), 1279–1296. <https://doi.org/10.1163/1568539042729685>
- Koda, H., Lemasson, A., Rizaldi, Oyakawa, C., Pamungkas, J., & Masataka, N. (2013). Possible Role of Mother-Daughter Vocal Interactions on the Development of Species-Specific Song in Gibbons. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0071432>
- Koda, H., Oyakawa, C., Kato, A., & Masataka, N. (2007). Experimental Evidence for the Volitional Control of Vocal Production in an Immature. *Source: Behaviour*, *144*(6), 681–692.
- Kondo, N., Watanabe, S., & Izawa, E. I. (2010). A temporal rule in vocal exchange among large-billed crows *Corvus macrorhynchos* in Japan. *Ornithol. Sci.*, *9*, 83–91. <https://doi.org/doi:10.2326/osj.9.83>
- Krause, M. A., Udell, M. A. R., Leavens, D. A., & Skopos, L. (2018). Animal pointing: Changing trends and findings from 30 years of research. *Journal of Comparative Psychology*, *132*(3), 326–345. <https://doi.org/10.1037/com0000125>
- Kroodsma, D. E. (1977). Correlates of Song Organization Among North American Wrens. *The American Naturalist*, *111*(981), 995–1008.
- Kuhl, P. K., & Meltzoff, A. N. (1982). The bimodal perception of speech in infancy. *Science*, *218*, 1138–1141.
- Kuhlwilm, M., de Manuel, M., Nater, A., Greminger, M. P., Krützen, M., & Marques-Bonet, T. (2016). Evolution and demography of the great apes. *Current Opinion in Genetics and Development*, *41*, 124–129. <https://doi.org/10.1016/j.gde.2016.09.005>
- Kulahci, I. G., Rubenstein, D. I., & Ghazanfar, A. A. (2015). Lemurs groom-at-a-distance through vocal networks. *Animal Behaviour*, *110*, 179–185. <https://doi.org/10.1016/j.anbehav.2015.09.016>

- Kumar, S., Stecher, G., Suleski, M., & Hedges, S. B. (2017). TimeTree: a resource for timelines, timetrees, and divergence times. *Molecular Biology and Evolution*, *34*, 1812–1819.
- Kummer, H. (1971). *Primate Societies: Group Techniques of Ecological Adaptation*. Chicago: Aldine.
- Kurtić, E., Brown, G. J., & Wells, B. (2013). Resources for turn competition in overlapping talk. *Speech Communication*, *55*, 1–23. <https://doi.org/10.1016/j.specom.2012.10.002>
- Ladefoged, P., & Disner, S. (2012). *Vowels and consonants 3rd edn*. Oxford, UK: Wiley-Blackwell.
- Lameira, A. R., Hardus, M. E., Bartlett, A. M., Shumaker, R. W., Wich, S. A., & Menken, S. B. J. (2015). Speech-like rhythm in a voiced and voiceless orangutan call. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0116136>
- Lameira, A. R., Hardus, M. E., Kowalsky, B., de Vries, H., Spruijt, B. M., Sterck, E. H. M., et al. (2013). Orangutan (*Pongo* spp.) whistling and implications for the emergence of an open-ended call repertoire: A replication and extension. *The Journal of the Acoustical Society of America*, *134*(3), 2326–2335. <https://doi.org/10.1121/1.4817929>
- Langmore, N. E. (2002). Vocal duetting: definitions, discoveries and directions. *Trends in Ecology and Evolution*, *17*(10), 451–452.
- Laporte, M. N. C., & Zuberbühler, K. (2010). Vocal greeting behaviour in wild chimpanzee females. *Animal Behaviour*, *80*(3), 467–473. <https://doi.org/10.1016/j.anbehav.2010.06.005>
- Laporte, M. N. C., & Zuberbühler, K. (2011). The development of a greeting signal in wild chimpanzees. *Developmental Science*, *14*(5), 1220–1234. <https://doi.org/10.1111/j.1467-7687.2011.01069.x>
- Leavens, D. A., Russell, J. L., & Hopkins, W. D. (2005). Intentionality as Measured in the Persistence and Elaboration of Communication by Chimpanzees (*Pan troglodytes*). *Child Development*, *76*, 291–306.
- Lehmann, J., & Boesch, C. (2003). Social influences on ranging patterns among chimpanzees (*Pan troglodytes verus*) in the Taï National Park, Côte d’Ivoire. *Behavioral Ecology*, *14*(5), 642–649. <https://doi.org/10.1093/beheco/arg047>
- Lehmann, J., & Boesch, C. (2008). Sexual differences in chimpanzee sociality. *International Journal of Primatology*, *29*(1), 65–81. <https://doi.org/10.1007/s10764-007-9230-9>
- Lehmann, J., & Boesch, C. (2009a). Sociality of the dispersing sex: the nature of social bonds in West African female chimpanzees, *Pan troglodytes*. *Animal Behaviour*, *77*, 377–387. <https://doi.org/10.1016/j.anbehav.2008.09.038>
- Lehmann, J., & Dunbar, R. I. M. (2009b). Network cohesion, group size and neocortex size in female-bonded Old World primates. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1677), 4417–4422. <https://doi.org/10.1098/rspb.2009.1409>
- Leigh, S. R., & Shea, B. T. (1995). Ontogeny and the Evolution of Adult Body Size Dimorphism in Apes. *American Journal of Primatology*, *36*, 37–60.
- Lemasson, A. (2011). What can forest guenons « tell » us about the origin of language? In A. Vilain, A., Schwartz, J. L. & J. C., Vauclair (Eds.), *Primate Communication and Human Language: Vocalisation, gestures, imitation and deixis in humans and non-humans* (pp. 39–70). Amsterdam: John Benjamins Publishing Company.
- Lemasson, A., Blois-Heulin, C., Jubin, R., & Hausberger, M. (2006). Female social relationships in a captive group of Campbell’s monkeys (*Cercopithecus campbelli campbelli*). *American Journal of Primatology*, *68*, 1161–1170. <https://doi.org/10.1002/ajp.20315>
- Lemasson, A., Gandon, E., & Hausberger, M. (2010). Attention to elders’ voice in non-human

- primates. *Biology Letters*, 6(3), 325–328. <https://doi.org/10.1098/rsbl.2009.0875>
- Lemasson, A., Gautier, J. P., & Hausberger, M. (2003). Vocal similarities and social bonds Campbell's monkey (*Cercopithecus campbelli*). *Comptes Rendus - Biologies*, 326(12), 1185–1193. <https://doi.org/10.1016/j.crvi.2003.10.005>
- Lemasson, A., Glas, L., Barbu, S., Lacroix, A., Guilloux, M., Remeuf, K., & Koda, H. (2011). Youngsters do not pay attention to conversational rules: Is this so for nonhuman primates? *Scientific Reports*, 1, 12–15. <https://doi.org/10.1038/srep00022>
- Lemasson, A., Guilloux, M., Rizaldi, Barbu, S., Lacroix, A., & Koda, H. (2013a). Age- and sex-dependent contact call usage in Japanese macaques. *Primates*, 54, 283–291. <https://doi.org/10.1007/s10329-013-0347-5>
- Lemasson, A., & Hausberger, M. (2004). Patterns of vocal sharing and social dynamics in a captive group of Campbell's monkeys (*Cercopithecus campbelli campbelli*). *Journal of Comparative Psychology*, 118(3), 347–359. <https://doi.org/10.1037/0735-7036.118.3.347>
- Lemasson, A., & Hausberger, M. (2011). Acoustic variability and social significance of calls in female Campbell's monkeys (*Cercopithecus campbelli campbelli*). *The Journal of the Acoustical Society of America*, 129(5), 3341–3352. <https://doi.org/10.1121/1.3569704>
- Lemasson, A., Jubin, R., Masataka, N., & Arlet, M. (2016). Copying hierarchical leaders' voices? Acoustic plasticity in female Japanese macaques. *Scientific Reports*, 6(21289). <https://doi.org/10.1038/srep21289>
- Lemasson A., Ouattara, K., & Zuberbühler, K. (2013b). Exploring the gaps between primate calls and human language. In M. Botha, R. & Everaert (Ed.), *The evolutionary emergence of language: evidence and inference* (pp. 181–203). Utrecht: Oxford University Press.
- Lemasson, A., Pereira, H., & Levréro, F. (2018). Social Basis of Vocal Interactions in Western Lowland Gorillas (*Gorilla g. gorilla*). *Journal of Comparative Psychology*, 132(2), 141–151. <https://doi.org/10.1037/com0000105>
- Lemasson, A., Remeuf, K., Rossard, A., & Zimmermann, E. (2012). Cross-Taxa Similarities in Affect-Induced Changes of Vocal Behavior and Voice in Arboreal Monkeys. *PLoS ONE*, 7, e45106. <https://doi.org/10.1371/journal.pone.0045106>
- Leroux, M., & Townsend, S. W. (2020). Call combinations in great apes and the evolution of syntax. *Animal Behavior and Cognition*, 7(2), 131–139. <https://doi.org/10.26451/abc.07.02.07.2020>
- Levé, M., Sueur, C., Petit, O., Matsuzawa, T., & Hirata, S. (2016). Social grooming network in captive chimpanzees: does the wild or captive origin of group members affect sociality? *Primates*, 57, 73–82. <https://doi.org/10.1007/s10329-015-0494-y>
- Levinson, S. C. (2006). On the Human “Interaction Engine.” In N.J. Enfield and S.C. Levinson (Ed.), *Roots of Human Sociality* (pp. 39–69). New York: Berg Publishers.
- Levinson, S. C. (2016). Turn-taking in Human Communication – Origins and Implications for Language Processing. *Trends in Cognitive Sciences*, 20(1), 6–14. <https://doi.org/10.1016/j.tics.2015.10.010>
- Levinson, S. C., & Holler, J. (2014). The origin of human multi-modal communication. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1651). <https://doi.org/10.1098/rstb.2013.0302>
- Levréro, F., Carrete-Vega, G., Herbert, A., Lawabi, I., Courtiol, A., Willaume, E., et al. (2015). Social shaping of voices does not impair phenotype matching of kinship in mandrills. *Nature Communications*, 6, 7609. <https://doi.org/10.1038/ncomms8609>
- Levréro, F., Clerc, O., Touitou, S., Fredet, J., Guéry, J.-P., & Lemasson, A. (2017). Bonobos converse

- and pay attention to breaking conversational rules. In S. e da V. (ISPA); A. for the S. of A. B. (ASAB) Instituto Universitário de Ciências Psicológicas (Ed.), *Behaviour 2017 - 35th International Ethological Conference* (p. hal-01576988). Estoril, Portugal.
- Lévréro, F., Gatti, S., Ménard, N., Petit, E., Caillaud, D., & Gautier-Hion, A. (2006). Living in nonbreeding groups: An alternative strategy for maturing gorillas. *American Journal of Primatology*, 68(3), 275–291. <https://doi.org/10.1002/ajp.20223>
- Lévréro, F., & Mathevon, N. (2013). Vocal Signature in Wild Infant Chimpanzees. *American Journal of Primatology*, 75(4), 324–332. <https://doi.org/10.1002/ajp.22108>
- Lévréro, F., Touitou, S., Frédet, J., Nairaud, B., Guéry, J.-P., & Lemasson, A. (2019). Social bonding drives vocal exchanges in Bonobos. *Scientific Reports*, 9, 711. <https://doi.org/10.1038/s41598-018-36024-9>
- Lewis, M., & Goldberg, S. (1969). The Acquisition and Violation of Expectancy: An Experimental Paradigm. *Journal of Experimental Child Psychology*, 7(1), 70–80.
- Liebal, K., Pika, S. S., & Tomasello, M. (2006). Gestural communication of orangutans (*Pongo pygmaeus*). *Gesture*, 6(1), 1–38.
- Liebal, K., Waller, B. M., Burrows, A. M., & Slocombe, K. E. (2013). *Primate Communication: A Multimodal Approach*. Cambridge: Cambridge University Press.
- Liebal, K., Waller, B. M., Burrows, A. M., & Slocombe, K. E. (2014a). Intentionality. In K. E. Liebal, K., Waller, B. M., Burrows, A.M., Slocombe (Ed.), *In Primate communication: a multimodal approach* (pp. 169–193). Cambridge: Cambridge University Press.
- Liebal, K., Waller, B. M., Burrows, A. M., & Slocombe, K. E. (2014b). *Primate Communication: A Multimodal Approach*. Cambridge: Cambridge University Press.
- Liebal, K., Waller, B. M., Burrows, A. M., & Slocombe K.E. (2012). *Primate communication: a multimodal approach*. Cambridge: Cambridge University Press.
- Lieberman, P. (2007). The Evolution of Human Speech Its Anatomical and Neural Bases. *Current Anthropology*, 48(1), 39–66. <https://doi.org/10.1086/509092>
- Lindsay, L., Gambi, C., & Rabagliati, H. (2019). Preschoolers Optimize the Timing of Their Conversational Turns Through Flexible Coordination of Language Comprehension and Production. *Psychological Science*, 30(4), 504–515.
- Locke, D. P., Hillier, L. W., Warren, W. C., Worley, K. C., Nazareth, L. V., Muzny, D. M., et al. (2011). Comparative and demographic analysis of orang-utan genomes. *Nature*, 469(7331), 529–533. <https://doi.org/10.1038/nature09687>
- Locke, J. L., & Snow, C. (1997). Social influences on vocal learning in human and nonhuman primates. In M. Snowdon, C.T.c Hausberger (Ed.), *Social influences on vocal development* (pp. 274–292). Cambridge, UK: Cambridge University Press.
- Lou-Magnuson, M., & Onnis, L. (2018). Social Network Limits Language Complexity. *Cognitive Science*, 42(8), 2790–2817. <https://doi.org/10.1111/cogs.12683>
- Luef, E. M., Breuer, T., & Pika, S. (2016). Food-associated calling in gorillas (*Gorilla g. gorilla*) in the wild. *PLoS ONE*, 11(2). <https://doi.org/10.1371/journal.pone.0144197>
- MacDonald, E. N., Johnson, E. K., Forsythe, J., Plante, P., & Munhall, K. G. (2012). Children’s development of self-regulation in speech production. *Current Biology*, 22, 113–117. <https://doi.org/10.1016/j.cub.2011.11.052>
- Mackinnon, J. (1974). The behaviour and ecology of wild orang-utans (*Pongo pygmaeus*). *Animal Behaviour*, 22, 3–74.



- MacLean, E. L., Merritt, D. J., & Brannon, E. M. (2008). Social complexity predicts transitive reasoning in prosimian primates. *Animal Behaviour*, *76*(2), 479–486. <https://doi.org/10.1016/j.anbehav.2008.01.025>
- Magliocca, F., & Gautier-Hion, A. (2002). Mineral content as a basis for food selection by western lowland gorillas in a forest clearing. *American Journal of Primatology*, *57*(2), 67–77. <https://doi.org/10.1002/ajp.10034>
- Mailund, T., Halager, A. E., Westergaard, M., Duthiel, J. Y., Munch, K., Andersen, L. N., et al. (2012). A New Isolation with Migration Model along Complete Genomes Infers Very Different Divergence Processes among Closely Related Great Ape Species. *PLoS Genetics*, *8*(12), e1003125. <https://doi.org/10.1371/journal.pgen.1003125>
- Mailund, T., Munch, K., & Schierup, M. H. (2014). Lineage Sorting in Apes. *Annual Review of Genetics*, *48*(1), 519–535. <https://doi.org/10.1146/annurev-genet-120213-092532>
- Maisels, F., Strindberg, S., Breuer, T., Greer, D., Jeffery, K., & Stokes, E. (2018). *Gorilla gorilla ssp. gorilla (amended version of 2016 assessment)*. *The IUCN Red List of Threatened Species 2018*.
- Maninger, N., Mendoza, S. P., Williams, D. R., Mason, W. A., Cherry, S. R., Rowland, D. J., et al. (2017). Imaging, behavior and endocrine analysis of “jealousy” in a monogamous primate. *Frontiers in Ecology and Evolution*, *5*(119). <https://doi.org/10.3389/fevo.2017.00119>
- Manser, M. B., Jansen, D. A. W. A. M., Graw, B., Hollén, L. I., Bousquet, C. A. H., Furrer, R. D., & le Roux, A. (2014). Vocal Complexity in Meerkats and Other Mongoose Species. In *Advances in the Study of Behavior* (Vol. 46, pp. 281–310). Academic Press Inc. <https://doi.org/10.1016/B978-0-12-800286-5.00006-7>
- Margulis, S. W., Whitham, J. C., & Ogorzalek, K. (2003). Silverback male presence and group stability in gorillas (*Gorilla gorilla gorilla*). *Folia Primatologica*, *74*(2), 92–96.
- Marler, P. (1965). Communication in monkeys and apes. In I. DeVore (Ed.), *Primate behavior. Field studies of monkeys and apes* (pp. 544–584). New York: Holt, Rinehart and Winston Inc.
- Marler, P. (1976). Social organization, communication and graded signals: The chimpanzee and the gorilla. In P. P. G. Bateson & R. A. Hinde (Ed.), *Growing points in ethology*. Cambridge, UK: Cambridge University Press.
- Marler, P., Mundinger, P., Sue Waser, M., & Lutjen, A. (1972). Effects of acoustical stimulation and deprivation on song development in red-winged blackbirds (*Agelaius phoeniceus*). *Animal Behaviour*, *20*, 586–606.
- Marler, P., & Tenaza, R. (1977). Signaling behavior of apes with special reference to vocalizations. In T. A. Sebeok (Ed.), *How Animals Communicate* (pp. 965–1033). Bloomington: Indiana University Press.
- Martinez, G., & Garcia, C. (2020). Sexual selection and sperm diversity in primates. *Molecular and Cellular Endocrinology*, *518*(110974), undefined-. <https://doi.org/10.1016/j.mce.2020.110974>
- Masataka, N. (1985). Development of Vocal Recognition of Mothers in Infant Japanese Macaques. *Developmental Psychobiology*, *18*(2), 107–114.
- Masataka, N. (2003). *The onset of language. The Onset of Language*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511489754>
- Masataka, N. (2008). *The origins of language: Unraveling evolutionary forces. The Origins of Language: Unraveling Evolutionary Forces*. <https://doi.org/10.1007/978-4-431-79102-7>
- Masataka, N., & Biben, M. (1987). Temporal Rules Regulating Affiliative Vocal Exchanges of Squirrel Monkeys. *Behaviour*, *101*, 311–319.

- Mattle-Greminger, M. P. (2015). *Unraveling the evolutionary history of Orangutans (genus: Pongo)- the impact of environmental processes and the genomic basis of adaptation*. University of Zurich. Retrieved from [www.zora.uzh.ch/urn:doi:10.5167/uzh-121397](http://www.zora.uzh.ch/urn:doi:10.5167/uzh-121397)
- Mazzini, F., Townsend, S. W., Virányi, Z., & Range, F. (2013). Wolf howling is mediated by relationship quality rather than underlying emotional stress. *Current Biology*, 23(17), 1677–1680. <https://doi.org/10.1016/j.cub.2013.06.066>
- McComb, K., & Semple, S. (2005). Coevolution of vocal communication and sociality in primates. *Biology Letters*, 1(4), 381–385. <https://doi.org/10.1098/rsbl.2005.0366>
- McCowan, B., & Reiss, D. (1997). Vocal learning in captive bottlenose dolphins: a comparison with humans and nonhuman animals. In C. T. Snodwon & M. Hausberger (Eds.), *Social Influences on Vocal Development* (pp. 178–207). Cambridge, UK: Cambridge University Press.
- McDaniel, R. R. J., & Driebe, D. J. (2005). *Uncertainty and surprise in complex systems: questions on working with the unexpected*. Berlin: Springer.
- McFarland, R., & Majolo, B. (2011). Exploring the components, asymmetry and distribution of relationship quality in wild barbary macaques (*macaca sylvanus*). *PLoS ONE*, 6(12), e28826. <https://doi.org/10.1371/journal.pone.0028826>
- McManus, K. F., Kelley, J. L., Song, S., Veeramah, K. R., Woerner, A. E., Stevison, L. S., et al. (2015). Inference of gorilla demographic and selective history from whole-genome sequence data. *Molecular Biology and Evolution*, 32(3), 600–612. <https://doi.org/10.1093/molbev/msu394>
- McNeil, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago: Chicago Press.
- McShea, D. W. (1991). Complexity and Evolution: What Everybody Knows. *Biology and Philosophy*, 6, 303–324.
- Meguerditchian, A., Cochet, H., & Vauclair, J. (2011a). From gesture to language. In V. J. Vilain A., Schwartz J.-L., Abry C. (Ed.), *Primate communication and human language: Vocalisation, gestures, imitation and deixis in humans and nonhumans* (pp. 91–118). Amsterdam: John Benjamins.
- Meguerditchian, A., Molesti, S., & Vauclair, J. (2011b). Right-handedness predominance in 162 baboons (*Papio anubis*) for gestural communication: Consistency across time and groups. *Behavioral Neuroscience*, 125(4), 653–660. <https://doi.org/10.1037/a0023823>
- Meguerditchian, A., & Vauclair, J. (2008). Vocal and gestural communication in nonhuman primates and the question of the origin of language. In L. Roska-Hardy & E.M. Neumann-Held (Ed.), *Learning from animals? Examining the Nature of Human Uniqueness* (Psychology Press., pp. 61–85). London. <https://doi.org/10.4324/9780203889909>
- Meguerditchian, A., & Vauclair, J. (2014). Communicative Signaling, Lateralization and Brain Substrate in Nonhuman Primates: Toward a Gestural or a Multimodal Origin of Language? *Mente Journal of Philosophical Studies*, 27, 135–160.
- Meltzoff, A. N., & Borton, R. W. (1979). Intermodal matching by human neonates. *Nature*, 282, 403–404.
- Mennill, D. J., & Ratcliffe, L. M. (2004). Overlapping and matching in the song contests of black-capped chickadees. *Animal Behaviour*, 67, 441–450. <https://doi.org/10.1016/j.anbehav.2003.04.010>
- Meunier, H., Fizet, J., & Vauclair, J. (2013). Tonkean macaques communicate with their right hand. *Brain and Language*, 126(2), 181–187. <https://doi.org/10.1016/j.bandl.2013.05.004>
- Miller, C. T., Flusberg, S., & Hauser, M. D. (2003). Interruption of long call production in tamarins: implications for vocal control. *Journal of Experimental Biology*, 206, 2629–2639.

- <https://doi.org/10.1242/jeb.00458>
- Miller, C. T., Mandel, K., & Wang, X. (2010). The communicative content of the common marmoset phee call during antiphonal calling. *American Journal of Primatology*, 72(11), 974–980. <https://doi.org/10.1002/ajp.20854>
- Miller, P. J. O., Shapiro, A. D., Tyack, P. L., & Solow, A. R. (2004). Call-type matching in vocal exchanges of free-ranging resident killer whales, *Orcinus orca*. *Animal Behaviour*, 67(6), 1099–1107. <https://doi.org/10.1016/j.anbehav.2003.06.017>
- Mitani, J. C. (1985). Sexual selection and adult male orangutan long calls. *Animal Behaviour*, 33(272–283), 272–283.
- Mitani, J. C. (1990). Experimental Field Studies of Asian Ape Social Systems. *International Journal of Primatology*, 11(2), 103–126.
- Mitani, J. C. (2009a). Male chimpanzees form enduring and equitable social bonds. *Animal Behaviour*, 77(3), 633–640. <https://doi.org/10.1016/j.anbehav.2008.11.021>
- Mitani, J. C. (2009b). Cooperation and competition in chimpanzees: Current understanding and future challenges. *Evolutionary Anthropology*, 18, 215–227. <https://doi.org/10.1002/evan.20229>
- Mitani, J. C., & Brandt, K. L. (1994). Social Factors Influence the Acoustic Variability in the Long-distance Calls of Male Chimpanzees. *Ethology*, 96(3), 233–252. <https://doi.org/10.1111/j.1439-0310.1994.tb01012.x>
- Mitani, J. C., & Gros-Louis, J. (1998). Chorusing and call convergence in chimpanzees: tests of three hypotheses. *Behaviour*, 135(9–10), 1041–1064.
- Mitani, J. C., & Watts, D. P. (2001). Why do chimpanzees hunt and share meat? *Animal Behaviour*, 61, 915–924. <https://doi.org/10.1006/anbe.2000.1681>
- Mitani, J. C., Watts, D. P., & Lwanga, J. S. (2002). Ecological and social correlates of chimpanzee party size and composition. In L. M. C. Boesch, G. Hohmann (Ed.), *Behavioural diversity in chimpanzees and bonobos* (pp. 102–111). Cambridge: Cambridge University Press.
- Morisaka, T., Yoshida, Y., Akune, Y., Mishima, H., & Nishimoto, S. (2013). Exchange of “signature” calls in captive belugas (*Delphinapterus leucas*). *Journal of Ethology*, 31(2), 141–149. <https://doi.org/10.1007/s10164-013-0358-0>
- Morrill, R. J., Paukner, A., Ferrari, P. F., & Ghazanfar, A. A. (2012). Monkey lipsmacking develops like the human speech rhythm. *Developmental Science*, 15(4), 1–12. <https://doi.org/10.1111/j.1467-7687.2012.01149.x>
- Morrogh-Bernard, H. C., Morf, N. V., Chivers, D. J., & Krützen, M. (2011). Dispersal Patterns of Orang-utans (*Pongo* spp.) in a Bornean Peat-swamp Forest. *International Journal of Primatology*, 32(2), 362–376. <https://doi.org/10.1007/s10764-010-9474-7>
- Moscovice, L. R., Deschner, T., & Hohmann, G. (2015). Welcome back: Responses of female bonobos (*Pan paniscus*) to fusions. *PLoS ONE*, 10(5), 1–16. <https://doi.org/10.1371/journal.pone.0127305>
- Mulavwa, M., Furuichi, T., Yangozene, K., Yamba-Yamba, M., Motema-Salo, B., Idani, ichi, et al. (2008). Seasonal Changes in Fruit Production and Party Size of Bonobos at Wamba. In J. Furuichi, T., & Thompson (Ed.), *The bonobos: behavior, ecology, and conservation* (pp. 121–134). New York: Springer.
- Müller, A. E., & Anzenberger, G. (2002). Duetting in the titi monkey *Callicebus cupreus*: Structure, pair specificity and development of duets. *Folia Primatologica*, 104–115. <https://doi.org/10.1159/000064788>

- Müller, M. N., & Mitani, J. C. (2005). Conflict and Cooperation in Wild Chimpanzees. *Advances in the Study of Behavior*, 35, 275–331. [https://doi.org/10.1016/S0065-3454\(05\)35007-8](https://doi.org/10.1016/S0065-3454(05)35007-8)
- Munhall, K., & Vatikiotis-Bateson, E. (1998). The moving face during speech communication. In R. Campbell, B. Dodd, & D. Burnham (Eds.), *Hearing by Eye II* (pp. 123–139). Sussex: Taylor and Francis.
- Murray, C. M., Eberly, L. E., & Pusey, A. E. (2006). Foraging strategies as a function of season and rank among wild female chimpanzees (*Pan troglodytes*). *Behavioral Ecology*, 17(6), 1020–1028. <https://doi.org/10.1093/beheco/arl042>
- Naguib, M., & Kipper, S. (2006). Effects of different levels of song overlapping on singing behaviour in male territorial nightingales (*Luscinia megarhynchos*). *Behavioral Ecology and Sociobiology*, 59(3), 419–426. <https://doi.org/10.1007/s00265-005-0066-z>
- Naguib, M., & Mennill, D. J. (2010). The signal value of birdsong: empirical evidence suggests song overlapping is a signal. *Animal Behaviour*, 80, 11–15. <https://doi.org/10.1016/j.anbehav.2010.06.001>
- Naguib, M., & Wiley, R. H. (2001). Estimating the distance to a source of sound: Mechanisms and adaptations for long-range communication. *Animal Behaviour*, 62(5), 825–837. <https://doi.org/10.1006/anbe.2001.1860>
- Nakahara, F., & Miyazaki, N. (2011). Vocal exchanges of signature whistles in bottlenose dolphins (*Tursiops truncatus*). *Journal of Ethology*, 29(2), 309–320. <https://doi.org/10.1007/s10164-010-0259-4>
- Nater, A., Greminger, M. P., Arora, N., van Schaik, C. P., Goossens, B., Singleton, I., et al. (2015). Reconstructing the demographic history of orang-utans using Approximate Bayesian Computation. *Molecular Ecology*, 24(2), 310–327. <https://doi.org/10.1111/mec.13027>
- Nater, A., Mattle-Greminger, M. P., Nurcahyo, A., Nowak, M. G., de Manuel, M., Desai, T., et al. (2017). Morphometric, Behavioral, and Genomic Evidence for a New Orangutan Species. *Current Biology*, 27(22), 3487–3498.e10. <https://doi.org/10.1016/j.cub.2017.09.047>
- Nater, A., Nietlisbach, P., Arora, N., van Schaik, C. P., van Noordwijk, M. A., Willems, E. P., et al. (2011). Sex-biased dispersal and volcanic activities shaped phylogeographic patterns of extant orangutans (genus: *Pongo*). *Molecular Biology and Evolution*, 28(8), 2275–2288. <https://doi.org/10.1093/molbev/msr042>
- Newman, J. D., & Symmes, D. (1974). Vocal Pathology in Socially Deprived Monkeys. *Developmental Psychobiology*, 7, 351–358.
- Newman, J. D., & Symmes, D. (1982). Inheritance and experience in the acquisition of primate acoustic behavior. In *Primate communication* (pp. 259–278). Cambridge, UK: Cambridge University Press.
- Newton-Fisher, N. E. (2002a). Relationships of male chimpanzee in the Budongo Forest, Uganda. In G. H. & L. M. C. Boesch (Ed.), *Behavioural Diversity in Chimpanzees and Bonobos* (pp. 125–137). Cambridge, UK: Cambridge University Press.
- Newton-Fisher, N. E. (2002b). Ranging patterns of male chimpanzees in the Budongo forest, Uganda: Range structure and individual differences. In & B. R. S. C. S. Harcourt (Ed.), *New perspectives in primate evolution and behaviour* (pp. 287–308). Otley: Westbury Scientific Publishing.
- Newton-Fisher, N. E., Thompson, M. E., Reynolds, V., Boesch, C., & Vigilant, L. (2010). Paternity and social rank in wild chimpanzees (*Pan troglodytes*) from the Budongo Forest, Uganda. *American Journal of Physical Anthropology*, 142(3), 417–428. <https://doi.org/10.1002/ajpa.21241>

- Nishida, T. (1968). The social group of wild chimpanzees in the Mahali Mountains. *Primates*, 9, 167–224. <https://doi.org/10.1007/BF01730971>
- Nishida, T. (1969). The social group of wild chimpanzees in the Mahale Mountains. *Primates*, 9, 167–224.
- Nishida, T. (1979). The social structure of chimpanzees of the Mahale Mountains. In D. Hamburg & E. McCown (Eds.), *The Great Apes* (pp. 72–121). Menlo Park, CA: Benjamin Cummings.
- Nishida, T. (1983). Alpha status and agonistic alliance in wild chimpanzees (*Pan troglodytes schweinfurthii*). *Primates*, 24(3), 318–336. <https://doi.org/10.1007/BF02381978>
- Nishida, T. (1990). *The Chimpanzees of the Mahale Mountains: Sexual and Life History Strategies*. Tokyo: University of Tokyo Press.
- Nishida, T., & Hosaka, K. (1996). Coalition strategies among adult male chimpanzees of the Mahale Mountains, Tanzania. In & T. N. W. McGrew, L. Marchant (Ed.), *Great Ape Societies* (pp. 114–134). Cambridge: Cambridge Univ. Press.
- Nishitani, N., & Hari, R. (2000). Temporal dynamics of cortical representation for action. *Proceedings of the National Academy of Sciences of the United States of America*, 97, 913–918. [www.pnas.org](http://www.pnas.org)
- Noë, R., & Laporte, M. (2014). Socio-spatial cognition in vervet monkeys. *Animal Cognition*, 17(3), 597–607. <https://doi.org/10.1007/s10071-013-0690-3>
- Norcross, J. L., & Newman, J. D. (1993). Context and Gender-Specific Differences in the Acoustic Structure of Common Marmoset (*Callithrix jacchus*) Phee Calls. *American Journal of Primatology*, 30, 37–54.
- Notman, H., & Rendall, D. (2005). Contextual variation in chimpanzee pant hoots and its implications for referential communication. *Animal Behaviour*, 70(1), 177–190. <https://doi.org/10.1016/j.anbehav.2004.08.024>
- Nowak, L., & Highfield, R. (2011). *Supercooperators: Altruism, Evolution, and Why We Need Each Other to Succeed*. London: Canongate.
- Nowak, M. (1999). Walker’s Primates of the World. In Baltimore and London (Ed.), . John Hopkins University Press.
- Nummela, S. U., Jovanovic, V., de la Mothe, L., & Miller, C. T. (2017). Social Context-Dependent Activity in Marmoset Frontal Cortex Populations during Natural Conversations. *The Journal of Neuroscience*, 37(29), 7036–7047. <https://doi.org/10.1523/JNEUROSCI.0702-17.2017>
- Oates, J. F., Tutin, C. E. G., Humle, T., Wilson, M. L., Baillie, J. E. M., Balmforth, Z., et al. (2007). *Pan troglodytes: IUCN Red List of Threatened Species*.
- Ohala, J. (1975). Temporal Regulation of Speech. In G. Fant & M. A. A. Tatham (Eds.), *Auditory Analysis and Perception of Speech* (pp. 431–453). London: Academic Press.
- Okobi, D. E. J., Banerjee, A., Matheson, A. M. M., Phelps, S. M., & Long, M. A. (2019). Motor cortical control of vocal interaction in neotropical singing mice. *Science*, 363(6430), 979–983. <https://doi.org/10.1126/science.aau1758>
- Ord, T. J., & Stamps, J. A. (2008). Alert signals enhance animal communication in “noisy” environments. *PNAS*, 105, 18830–18835. [www.pnas.org/cgi/doi/10.1073/pnas.0807657105](http://www.pnas.org/cgi/doi/10.1073/pnas.0807657105)
- Ouattara, K., Lemasson, A., & Zuberbühler, K. (2009a). Campbell’s monkeys concatenate vocalizations into context-specific call sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 106(51), 22026–22031. <https://doi.org/10.1073/pnas.0908118106>

- Ouattara, K., Zuberbühler, K., N'goran, E. K., Gombert, J. E., & Lemasson, A. (2009b). The alarm call system of female Campbell's monkeys. *Animal Behaviour*, 78(1), 35–44. <https://doi.org/10.1016/j.anbehav.2009.03.014>
- Owings, D. H., & Morton, E. S. (1998). *Animal vocal communication: a new approach*. Cambridge, UK: Cambridge University Press.
- Owren, M. J., Amoss, R. T., & Rendall, D. (2011). Two organizing principles of vocal production: Implications for nonhuman and human primates. *American Journal of Primatology*, 73, 530–544. <https://doi.org/10.1002/ajp.20913>
- Owren, M. J., Dieter, J. A., Seyfarth, R. M., & Cheney, D. L. (1992). “Food” Calls Produced by Adult Female Rhesus (*Macaca mulatta*) and Japanese (*M. fuscata*) Macaques, Their Normally-Raised Offspring, and Offspring Cross-Fostered between Species. *Behaviour*, 120(3), 218–231.
- Pagel, M. (2012). *Wired for Culture: The Natural History of Human Cooperation*. New York: Norton.
- Papworth, S., Böse, A.-S., Barker, J., Schel, A.-M., & Zuberbühler, K. (2008). Male blue monkeys alarm call in response to danger experienced by others. *Biology Letters*, 4, 472–475. <https://doi.org/10.1098/rsbl.2008.0299>
- Parish, A. R. (1994). Sex and Food Control in the “Uncommon Chimpanzee”: How Bonobo Females Overcome a Phylogenetic Legacy of Male Dominance. *Ethology and Sociobiology*, 15, 157–179.
- Parnell, R. J. (2002). Group size and structure in western lowland gorillas (*Gorilla gorilla gorilla*) at Mbeli Bai, Republic of Congo. *American Journal of Primatology*, 56(4), 193–206. <https://doi.org/10.1002/ajp.1074>
- Parrish, J. K., Viscido, S. V., & Grünbaum, D. (2002). Self-Organized Fish Schools: An Examination of Emergent Properties. *The biological Bulletin*, 202(3), 296–305. <https://www.jstor.org/stable/1543482>
- Patterson, F. G. (1978). The gestures of a gorilla: Language acquisition in another pongid. *Brain and Language*, 5(1), 72–97. [https://doi.org/10.1016/0093-934X\(78\)90008-1](https://doi.org/10.1016/0093-934X(78)90008-1)
- Patterson, N., Richter, D. J., Gnerre, S., Lander, E. S., & Reich, D. (2006). Genetic evidence for complex speciation of humans and chimpanzees. *Nature*, 441(7097), 1103–1108. <https://doi.org/10.1038/nature04789>
- Peckre, L., Kappeler, P. M., & Fichtel, C. (2019). Clarifying and expanding the social complexity hypothesis for communicative complexity. *Behavioral Ecology and Sociobiology*, 73(1). <https://doi.org/10.1007/s00265-018-2605-4>
- Pennec, F., Krief, S., Hladik, A., Ayingweu, C. L., Bortolamiol, S., Bokika Ngawolo, J. C., & Narat, V. (2016). Floristic and structural vegetation typology of bonobo habitats in a forest-Savanna Mosaic (Bolobo territory, D.R.Congo). *Plant Ecology and Evolution*, 149(2), 199–215. <https://doi.org/10.5091/plecevo.2016.1157>
- Pereira, A. S., Kavanagh, E., Hobaiter, C., Slocombe, K. E., & Lameira, A. R. (2020). Chimpanzee lip-smacks confirm primate continuity for speech-rhythm evolution. *Biology Letters*, 16(5), 20200232. <https://doi.org/10.1098/rsbl.2020.0232>
- Perry, S. (1996). Female-Female Social Relationships in Wild White-Faced Capuchin Monkeys, *Cebus capucinus*. *American Journal of Primatology*, 40, 40167–182.
- Pika, S. (2017). Unpeeling the layers of communicative complexity. *Animal Behaviour*, 134, 223–227. <https://doi.org/10.1016/j.anbehav.2017.07.002>
- Pika, S., Liebal, K., Call, J., & Tomasello, M. (2005). Gestural communication of apes. *Gesture*, 5(1/2), 41–56. <https://doi.org/10.1075/gest.5.1.05pik>

- Pika, S., Wilkinson, R., Kendrick, K. H., & Vernes, S. C. (2018). Taking turns: bridging the gap between human and animal communication. *Proceedings of the Royal Society B*, 285(1880), 1–9. <https://doi.org/10.1098/rspb.2018.0598>
- Pinker, S. (1994). *The language instinct. How the mind creates language*. New-York: Harper Perennial Modern Classics.
- Pistorio, A. L., Vintch, B., & Wang, X. (2006). Acoustic analysis of vocal development in a New World primate, the common marmoset ( *Callithrix jacchus* ). *The Journal of the Acoustical Society of America*, 120(3), 1655–1670. <https://doi.org/10.1121/1.2225899>
- Platzen, D., & Magrath, R. D. (2005). Adaptive differences in response to two types of parental alarm call in altricial nestlings. *Proceedings of the Royal Society B: Biological Sciences*, 272(1568), 1101–1106. <https://doi.org/10.1098/rspb.2005.3055>
- Ploog, D. (1981). Neurobiology of Primate Audio-Vocal Behavior. *Bruin Research Reviews*, 3(1), 35–61.
- Pollard, K. A., & Blumstein, D. T. (2012). Evolving communicative complexity: Insights from rodents and beyond. *Philosophical Transactions of the Royal Society B: Biological Sciences*. Royal Society. <https://doi.org/10.1098/rstb.2011.0221>
- Pollick, A. S., & de Waal, F. B. M. (2007). Ape gestures and language evolution. *Proceedings of the National Academy of Sciences*, 104(19), 8184–8189. [www.pnas.org/cgi/content/full/](http://www.pnas.org/cgi/content/full/)
- Prado-Martinez, J., Sudmant, P. H., Kidd, J. M., Li, H., Kelley, J. L., Lorente-Galdos, B., et al. (2013). Great ape genetic diversity and population history. *Nature*, 499(7459), 471–475. <https://doi.org/10.1038/nature12228>
- Price, P. H. (1979). Developmental Determinants of Structure in Zebra Finch Song. *Journal of Comparative and Physiological Psychology*, 93(2), 260–277.
- Prieur, J., Barbu, S., Blois-Heulin, C., & Lemasson, A. (2020). The origins of gestures and language: history, current advances and proposed theories. *Biological Reviews*, 95(3), 531–554. <https://doi.org/10.1111/brv.12576>
- Prieur, J., Pika, S., Barbu, S., & Blois-Heulin, C. (2016a). A multifactorial investigation of captive chimpanzees' intraspecific gestural laterality. *Animal Behaviour*, 116, 31–43. <https://doi.org/10.1016/j.anbehav.2016.03.024>
- Prieur, J., Pika, S., Barbu, S., & Blois-Heulin, C. (2016b). Gorillas are right-handed for their most frequent intraspecific gestures. *Animal Behaviour*, 118, 165–170. <https://doi.org/10.1016/j.anbehav.2016.06.008>
- Prüfer, K., Munch, K., Hellmann, I., Akagi, K., Miller, J. R., Walenz, B., et al. (2012). The bonobo genome compared with the chimpanzee and human genomes. *Nature*, 486(7404), 527–531. <https://doi.org/10.1038/nature11128>
- Pusey, A., Williams, J., & Goodall, J. (1997). The influence of dominance rank on the reproductive success of female chimpanzees. *Science*, 277(5327), 828–831. <https://doi.org/10.1126/science.277.5327.828>
- Quick, N. J., & Janik, V. M. (2012). Bottlenose dolphins exchange signature whistles when meeting at sea. *Proceedings of the Royal Society B: Biological Sciences*, 279(1738), 2539–2545. <https://doi.org/10.1098/rspb.2011.2537>
- Ramos-Fernández, G. (2005). Vocal communication in a fission-fusion society: Do spider monkeys stay in touch with close associates? *International Journal of Primatology*, 26(5), 1077–1092. <https://doi.org/10.1007/s10764-005-6459-z>
- Ravignani, A., Bowling, D., & Fitch, W. T. (2014). Chorusing, synchrony and the evolutionary

- functions of rhythm. *Frontiers in Psychology*, 5, 1118. <https://doi.org/10.3389/fpsyg.2014.01118>
- Reber, S. A., Šlipogor, V., Oh, J., Ravignani, A., Hoeschele, M., Bugnyar, T., & Fitch, W. T. (2019). Common marmosets are sensitive to simple dependencies at variable distances in an artificial grammar. *Evolution and Human Behavior*, 40(2), 214–221. <https://doi.org/10.1016/j.evolhumbehav.2018.11.006>
- Rebout, N., De Marco, A., Lone, J.-C., Sanna, A., Cozzolino, R., Micheletta, J., et al. (2020). Tolerant and intolerant macaques show different levels of structural complexity in their vocal communication. *Proceedings of the Royal Society B: Biological Sciences*, 287(1928), 20200439. <https://doi.org/10.1098/rspb.2020.0439>
- Reby, D., Cargnelutti, B., & Hewison, A. J. M. (1999). Contexts and possible functions of barking in roe deer. *Animal Behaviour*, 57, 1121–1128. <http://www.idealibrary.comon>
- Reddish, P., Bulbulia, J., & Fischer, R. (2014). Does synchrony promote generalized prosociality? *Religion, Brain and Behavior*, 4(1), 3–19.
- Reynolds, V. (2005). *The chimpanzees of the Budongo Forest*. Oxford: Oxford University Press.
- Rijksen, H. D. (1978). *field study on Sumatran Orang-Utans (Pongo pygmaeus abelii, lesson 1827): Ecology, Behaviour, and Conservation*. (W. Veenman, H. and B.V. Zonene, Ed.).
- Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, 21(5), 188–194.
- Robbins, M. M., Bermejo, M., Cipolletta, C., Magliocca, F., Parnell, R. J., & Stokes, E. (2004). Social structure and life-history patterns in western gorillas (*Gorilla gorilla gorilla*). *American Journal of Primatology*, 64, 145–159. <https://doi.org/10.1002/ajp.20069>
- Robinson, D. J., & Hall, M. J. (2002). Sound signalling in orthoptera. In P. Evans (Ed.), *Advances in Insect Physiology* (Vol. 29, pp. 151–278). Academic Press Inc. [https://doi.org/10.1016/s0065-2806\(02\)29003-7](https://doi.org/10.1016/s0065-2806(02)29003-7)
- Rodman, P. S. (1973). Population composition and adaptive organization among orangutans of the Kutai Reserve. In R. P. Crook, J.H. and Michael (Ed.), *Comparative Ecology and Behaviour of Primates* (pp. 171–209). London: Academic Press.
- Rodman, P. S. (1979). Individual activity patterns and the solitary nature of orangutans. In E. R. Hamburg, D.A. and McCown (Ed.), *The Great Apes* (pp. 235–255). Menlo Park, California: Benjamin/Cummings.
- Rodman, P. S. (1988). Diversity and consistency in ecology and behaviour. In J. H. Schwartz (Ed.), *Orang-Utan biology* (pp. 31–51). New York: Oxford University Press.
- Rodman, P. S., & Mitani, J. C. (1987). Orangutans: sexual dimorphism in a solitary species. In T. T. Smuts, B.B., Cheney, D.L., Seyfarth, R.M., Wrangham, R.M., and Struhsaker (Ed.), *Primate Societies* (pp. 146–154). Chicago: University of Chicago Press.
- Roeder, J.-J., & Gosset, D. L. (2001). Communication chez les primates non humains: de l'éthologie classique à l'éthologie cognitive. *Primatologie*, 4, 5–25.
- Roush, R. S., & Snowdon, C. T. (2001). Food Transfer and Development of Feeding Behavior and Food-Associated Vocalizations in Cotton-Top Tamarins. *Ethology*, 107(5), 415–429. [www.blackwell.de/synergy](http://www.blackwell.de/synergy)
- Rowe, N. (1996). *The Pictorial Guide to the Living Primates*. East Hampton, New York: Pogonias Press.
- Roy, S., Miller, C. T., Gottsch, D., & Wang, X. (2011). Vocal control by the common marmoset in the presence of interfering noise. *Journal of Experimental Biology*, 214, 3619–3629.



<https://doi.org/10.1242/jeb.056101>

- Rukstalis, M., & French, J. A. (2005). Vocal buffering of the stress response: Exposure to conspecific vocalizations moderates urinary cortisol excretion in isolated marmosets. *Hormones and Behavior*, 47(1), 1–7. <https://doi.org/10.1016/j.yhbeh.2004.09.004>
- Ruvolo, M., Disotell, T. R., Allardt, M. W., Brown, W. M., & Honeycutt, R. L. (1991). *Resolution of the African hominoid trichotomy by use of a mitochondrial gene sequence (hominoid phylogeny/mitochondrial DNA/cytochrome oxidase subunit H gene/molecular anthropology)*. *Proc. Natl. Acad. Sci. USA* (Vol. 88).
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, 50, 696–735. <https://doi.org/10.2307/412243>
- Salmi, R., Hammerschmidt, K., & Doran-Sheehy, D. M. (2013). Western gorilla vocal repertoire and contextual use of vocalizations. *Ethology*, 119(10), 831–847. <https://doi.org/10.1111/eth.12122>
- Salmi, R., Hammerschmidt, K., & Doran-Sheehy, D. M. (2014). Individual distinctiveness in call types of wild western female gorillas. *PLoS ONE*, 9(7), e101940. <https://doi.org/10.1371/journal.pone.0101940>
- Scally, A., Dutheil, J. Y., Hillier, L. W., Jordan, G. E., Goodhead, I., Herrero, J., et al. (2012). Insights into hominid evolution from the gorilla genome sequence. *Nature*, 483(7388), 169–175. <https://doi.org/10.1038/nature10842>
- Schaeffer, S. A. (2011). *Personality, subjective well-being, and vocal communication in captive male Western lowland gorillas*. The state University of New Jersey, New Jersey.
- Schamberg, I., Cheney, D. L., Clay, Z., Hohmann, G., & Seyfarth, R. M. (2016). Call combinations, vocal exchanges and interparty movement in wild bonobos. *Animal Behaviour*, 122, 109–116. <https://doi.org/10.1016/j.anbehav.2016.10.003>
- Schamberg, I., Cheney, D. L., Clay, Z., Hohmann, G., & Seyfarth, R. M. (2017). Bonobos use call combinations to facilitate inter-party travel recruitment. *Behavioral Ecology and Sociobiology*, 71, 75. <https://doi.org/10.1007/s00265-017-2301-9>
- Schel, A. M., Candiotti, A., & Zuberbühler, K. (2010). Predator-detering alarm call sequences in Guereza colobus monkeys are meaningful to conspecifics. *Animal Behaviour*, 80, 799–808. <https://doi.org/10.1016/j.anbehav.2010.07.012>
- Schel, A. M., Townsend, S. W., Machanda, Z., Zuberbühler, K., & Slocombe, K. E. (2013). Chimpanzee Alarm Call Production Meets Key Criteria for Intentionality. *PLoS ONE*, 8(10), e76674. <https://doi.org/10.1371/journal.pone.0076674>
- Schel, A. M., & Zuberbühler, K. (2012). Dawn chorusing in guereza colobus monkeys. *Behavioral Ecology and Sociobiology*, 66(3), 361–373. <https://doi.org/10.1007/s00265-011-1282-3>
- Schlenker, P., Chemla, E., Cäsar, C., Ryder, R., & Zuberbühler, K. (2017). Titi Semantics: Context and Meaning in Titi Monkey Call Sequences \*. *Nature Language & Linguistic Theory*, 35(1), 271–298.
- Schulz, T. M., Whitehead, H., Gero, S., & Rendell, L. (2008). Overlapping and matching of codas in vocal interactions between sperm whales: insights into communication function. *Animal Behaviour*, 76(6), 1977–1988. <https://doi.org/10.1016/j.anbehav.2008.07.032>
- Schürmann, C. L. (1982). Mating behaviour of wild orangutans. In L. E. M. de Boer (Ed.), *The Orang Utan: its biology and conservation* (pp. 269–284). The Hague: Junk.
- Schuster, P. (2016). How complexity originates: Examples from history reveal additional roots to complexity. *Complexity*, 21, 7–12. <https://doi.org/10.1002/cplx.21841>

- Schwabl, H., & Sonnenschein, E. (1992). Antiphonal duetting and sex hormones in the tropical bush shrike *Laniarius funebris*. *Hormones and Behavior*, 26, 295–307.
- Schwartz, J. J. (1987). The Function of Call Alternation in Anuran Amphibians: a Test of Three Hypothesis. *Evolution: International Journal of Organic Evolution*, 41(3), 461–471.
- Seddon, N., Butchart, S. H. M., & Odling-Smee, L. (2002). Duetting in the subdesert mesite *Monias benschi*: Evidence for acoustic mate defence? *Behavioral Ecology and Sociobiology*, 52(1), 7–16. <https://doi.org/10.1007/s00265-002-0488-9>
- Seibt, U., & Wickler, W. (1977). Duettieren als Revier-Anzeige bei Vogeln. *Zeitschrift für Tierpsychologie*, 43, 180–187.
- Serckx, A., Kühl, H. S., Beudels-Jamar, R. C., Poncin, P., Bastin, J. F., & Huynen, M. C. (2015). Feeding ecology of bonobos living in forest-savannah mosaics: Diet seasonal variation and importance of fallback foods. *American Journal of Primatology*, 77(9), 948–962. <https://doi.org/10.1002/ajp.22425>
- Sewall, K. B., Young, A. M., & Wright, T. F. (2016). Social calls provide novel insights into the evolution of vocal learning. *Animal Behaviour*, 120, 163–172. <https://doi.org/10.1016/j.anbehav.2016.07.031>
- Seyfarth, R. M. (1987). Vocal communication and its relation to language. *Primate societies*, 440–451.
- Seyfarth, R. M., & Cheney, D. L. (1986). Vocal development in vervet monkeys. *Animal Behaviour*, 34(6), 1640–1658. [https://doi.org/10.1016/S0003-3472\(86\)80252-4](https://doi.org/10.1016/S0003-3472(86)80252-4)
- Seyfarth, R. M., & Cheney, D. L. (1997). Some general features of vocal development in nonhuman primates. In C. T. Snowdon & M. Hausberger (Eds.), *Social Influences on Vocal Development* (pp. 249–273). Cambridge: Cambridge University Press.
- Seyfarth, R. M., & Cheney, D. L. (2003). Signalers and Receivers in Animal Communication. *Annual Review of Psychology*, 54, 145–173. <https://doi.org/10.1146/annurev.psych.54.101601.145121>
- Seyfarth, R. M., Cheney, D. L., Bergman, T., Fischer, J., Zuberbühler, K., & Hammerschmidt, K. (2010). The central importance of information in studies of animal communication. *Animal Behaviour*, 80(1), 3–8. <https://doi.org/10.1016/j.anbehav.2010.04.012>
- Seyfarth, R. M., Cheney, D. L., & Bergman, T. J. (2005). Primate social cognition and the origins of language. *Trends in Cognitive Sciences*, 9(6), 264–266. <https://doi.org/10.1016/j.tics.2005.04.001>
- Seyfarth, R. M., Cheney, D. L., Harcourt, A. H., & Stewart, K. J. (1994). The acoustic features of gorilla double grunts and their relation to behavior. *American Journal of Primatology*, 33, 31–50. <https://doi.org/10.1002/ajp.1350330104>
- Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervetmonkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour*, 28, 1070–1094.
- Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. *Science*, 210(4471), 801–803.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379–423.
- Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.
- Shepherd, S. V., & Freiwald, W. A. (2018). Functional Networks for Social Communication in the Macaque Monkey. *Neuron*, 99(2), 413–420.e3. <https://doi.org/10.1016/j.neuron.2018.06.027>
- Shultz, S., & Dunbar, R. I. M. (2006). Chimpanzee and felid diet composition is influenced by prey

- brain size. *Biology Letters*, 2(4), 505–508. <https://doi.org/10.1098/rsbl.2006.0519>
- Sicotte, P. (1993). Inter-Group Encounters and Female Transfer in Mountain Gorillas: Influence of Group Composition on Male Behavior. *American Journal of Primatology*, 30(1), 21–36.
- Siegel, S., & Castellan, N. J. (1998). *Nonparametric statistics for the behavioural sciences*. (Siegel & Castellan (eds), Eds.) (2nd Edition.). New-York: McGraw-Hill.
- Silk, J. B., Alberts, S. C., & Altmann, J. (2003). Social Bonds of Female Baboons Enhance Infant Survival. *Science*, 302(5648), 1231–1234. <https://doi.org/10.1126/science.1088474>
- Simões, C. S., Vianney, P. V. R., de Moura, M. M., Freire, M. A. M., Mello, L. E., Samesshima, K., et al. (2010). Activation of frontal neocortical areas by vocal production in marmosets. *Frontiers in Integrative Neuroscience*, 4, 1–12. <https://doi.org/10.3389/fnint.2010.00123>
- Slabbekoorn, H., & Peet, M. (2003). Birds sing at a higher pitch in urban noise. *Nature*, 424, 267-undefined.
- Slocombe, K. E., Waller, B. M., & Liebal, K. (2011). The language void: The need for multimodality in primate communication research. *Animal Behaviour*, 81, 919–924. <https://doi.org/10.1016/j.anbehav.2011.02.002>
- Slocombe, K. E., & Zuberbühler, K. (2005). Functionally referential communication in a chimpanzee. *Current Biology*, 15, 1779–1784. <https://doi.org/10.1016/j.cub.2005.08.068>
- Slocombe, K. E., & Zuberbühler, K. (2007). Chimpanzees modify recruitment screams as a function of audience composition. *Proceedings of the National Academy of Sciences of the United States of America*, 104(43), 17228–17233. <https://doi.org/10.1073/pnas.0706741104>
- Slocombe, K. E., & Zuberbühler, K. (2010). Vocal Communication in Chimpanzees. In *The Mind of the Chimpanzee: Ecological and Experimental Perspectives*. Chicago & London: University of Chicago Press.
- Smith, A., & Zelaznik, H. N. (2004). Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology*, 45(1), 22–33. <https://doi.org/10.1002/dev.20009>
- Smith, R. J., & Jungers, W. L. (1997). Body mass in comparative primatology. *Journal of Human Evolution*, 32, 523–559.
- Smith, W. J. (1965). Message, Meaning, and Context in Ethology. *American Naturalist*, 99(908), 405–409.
- Smith, W. J. (1969). Message of vertebrate communication. *Science*, 165, 145–158.
- Smith, W. J. (1994). Animal duets: forcing a mate to be attentive. *Journal of Theoretical Biology*, (166), 221–223.
- Snowdon, C. T. (2009). Plasticity of Communication in Nonhuman Primates. *Advances in the Study of Behavior*, 40, 239–276. [https://doi.org/10.1016/S0065-3454\(09\)40007-X](https://doi.org/10.1016/S0065-3454(09)40007-X)
- Snowdon, C. T., & Cleveland, J. (1984). “Conversations” among pygmy marmosets. *American Journal of Primatology*, 7(1), 15–20. <https://doi.org/10.1002/ajp.1350070104>
- Snowdon, C. T., & Cronin, K. A. (2007). Cooperative breeders do cooperate. *Behavioural Processes*, 76(2), 138–141.
- Snowdon, C. T., & Elowson, A. M. (1999). Pygmy marmosets modify call structure when paired. *Ethology*, 105(10), 893–908. <https://doi.org/10.1046/j.1439-0310.1999.00483.x>
- Snowdon, C. T., & Elowson, A. M. (2001). “Babbling” in Pygmy Marmosets: Development after Infancy. *Behaviour*, 138(10), 1235–1248.

- Snowdon, C. T., & Hausberger, M. (1997). *Social influences on vocal development*. Cambridge, UK: Cambridge University Press.
- Soltis, J., Bernhards, D., Donkin, H., & Newman, J. D. (2002). Squirrel monkey chuck call: Vocal response to playback chucks based on acoustic structure and affiliative relationship with the caller. *American Journal of Primatology*, 7, 119–130. <https://doi.org/10.1002/ajp.10039>
- Soltis, J., Leong, K., & Savage, A. (2005). African elephant vocal communication I: Antiphonal calling behaviour among affiliated females. *Animal Behaviour*, 70, 579–587. <https://doi.org/10.1016/j.anbehav.2004.11.015>
- Spehar, S. N., & Di Fiore, A. (2013). Loud calls as a mechanism of social coordination in a fission-fusion taxon, the white-bellied spider monkey (*Ateles belzebuth*). *Behavioral Ecology and Sociobiology*, 67(6), 947–961. <https://doi.org/10.1007/s00265-013-1520-y>
- Spillmann, B., Dunkel, L. P., van Noordwijk, M. A., Amda, R. N. A., Lameira, A. R., Wich, S. A., & van Schaik, C. P. (2010). Acoustic properties of long calls given by flanged male orang-utans (*Pongo pygmaeus wurmbii*) reflect both individual identity and context. *Ethology*, 116(5), 385–395. <https://doi.org/10.1111/j.1439-0310.2010.01744.x>
- Spillmann, B., Willems, E. P., van Noordwijk, M. A., Setia, T. M., & van Schaik, C. P. (2017). Confrontational assessment in the roving male promiscuity mating system of the Bornean orangutan. *Behavioral Ecology and Sociobiology*, 71(1), 20. <https://doi.org/10.1007/s00265-016-2252-6>
- Stephan, C., & Zuberbühler, K. (2008). Predation increases acoustic complexity in primate alarm calls. *Biology Letters*, 4(6), 641–644. <https://doi.org/10.1098/rsbl.2008.0488>
- Stevens, J. M. G., Vervaecke, H., de Vries, H., & van Elsacker, L. (2006). Social structures in pan paniscus: Testing the female bonding hypothesis. *Primates*, 47(3), 210–217. <https://doi.org/10.1007/s10329-005-0177-1>
- Stewart, K. J., & Harcourt, A. H. (1994). Gorillas' Vocalizations during Rest Periods: Signals of Impending Departure? *Source: Behaviour*, 130(1–2), 29–40. <http://www.jstor.org/stable/4535204>
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106(26), 10587–10592. <https://doi.org/10.1073/pnas.0903616106>
- Stoddard, P. K., & Beecher, M. D. (1983). Parental recognition of offspring in the cliff swallow. *Auk*, 100, 795–799.
- Stokes, E. J., Parnell, R. J., & Olejniczak, C. (2003). Female dispersal and reproductive success in wild western lowland gorillas (*Gorilla gorilla gorilla*). *Behavioral Ecology and Sociobiology*, 54(4), 329–339. <https://doi.org/10.1007/s00265-003-0630-3>
- Stone, E., & Trost, C. H. (1991). Predators, risks and context for mobbing and alarm calls in black-billed magpies. *Animal Behaviour*, 41, 633–638.
- Stumpf, R. M., & Boesch, C. (2005). Does promiscuous mating preclude female choice? Female sexual strategies in chimpanzees (*Pan troglodytes verus*) of the Taï National Park, Côte d'Ivoire. *Behavioral Ecology and Sociobiology*, 57(5), 511–524. <https://doi.org/10.1007/s00265-004-0868-4>
- Stumpner, A., & Meyer, S. (2001). Songs and the Function of Song Elements in Four Duetting Bushcricket Species (Ensifera, Phaneropteridae, Barbitistes). *Journal of Insect Behavior*, 14(4), 511–534.
- Sugardjito, J., & Nurhuda, N. (1981). Meat-eating Behaviour in Wild Orang Utans, *Pongo pygmaeus*. *PRIMATES*, 22(3), 414–416.

- Sugiura, H. (1993). Temporal and Acoustic Correlates in Vocal Exchange of Coo Calls in Japanese Macaques. *Behaviour*, 124, 3–4. <https://doi.org/10.1163/156853993X00588>
- Sugiura, H. (1998). Matching of acoustic features during the vocal exchange of coo calls by Japanese macaques. *Animal Behaviour*, 55, 673–687.
- Sugiura, H., & Masataka, N. (1995). Temporal and acoustic flexibility in vocal exchanges of coo calls in Japanese macaques (*Macaca fuscata*). In E. Zimmermann, J. D. Newman, & U. Jürgens (Eds.), *Current topics in primate vocal communication* (NY : Plenum Press., pp. 121–140). New York. <https://doi.org/10.1007/978-1-4757-9930-9>
- Surbeck, M., Girard-Buttoz, C., Boesch, C., Crockford, C., Fruth, B., Hohmann, G., et al. (2017). Sex-specific association patterns in bonobos and chimpanzees reflect species differences in cooperation. *Royal Society Open Science*, 4(5), 1–20. <https://doi.org/10.1098/rsos.161081>
- Surbeck, M., & Hohmann, G. (2008). Primate hunting by bonobos at LuiKotale, Salonga National Park. *Current Biology*, 18, 906–907.
- Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communications*, 7, 10986–10987. <https://doi.org/10.1038/ncomms10986>
- Taborsky, B., & Oliveira, R. F. (2012). Social competence: An evolutionary approach. *Trends in Ecology and Evolution*, 27(12), 679–688. <https://doi.org/10.1016/j.tree.2012.09.003>
- Tagliatalata, J. P., Russell, J. L., Schaeffer, J. A., & Hopkins, W. D. (2008). Communicative Signaling Activates “Broca’s” Homolog in Chimpanzees. *Current Biology*, 18(5), 343–348. <https://doi.org/10.1016/j.cub.2008.01.049>
- Tagliatalata, J. P., Russell, J. L., Schaeffer, J. A., & Hopkins, W. D. (2011). Chimpanzee vocal signaling points to a multimodal origin of human language. *PLoS ONE*, 6(4), e1885. <https://doi.org/10.1371/journal.pone.0018852>
- Takahashi, D. Y., Fenley, A. R., & Ghazanfar, A. A. (2016). Early development of turn-taking with parents shapes vocal acoustics in infant marmoset monkeys. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693), 20150370. <https://doi.org/10.1098/rstb.2015.0370>
- Takahashi, D. Y., Fenley, A. R., Teramoto, Y., Narayanan, D. Z., Borjon, J. I., Holmes, P., & Ghazanfar, A. A. (2015). The developmental dynamics of marmoset monkey vocal production. *Science*, 349(6249), 734–738. <https://doi.org/10.1126/science.aaa7945>
- Takahashi, D. Y., Liao, D. A., & Ghazanfar, A. A. (2017). Vocal Learning via Social Reinforcement by Infant Marmoset Monkeys. *Current Biology*, 27, 1844–1852. <https://doi.org/10.1016/j.cub.2017.05.004>
- Takahashi, D. Y., Narayanan, D., & Ghazanfar, A. A. (2012). A computational model for vocal exchange dynamics and their development in marmoset monkeys. In *International Conference on Development and Learning and Epigenetic Robotics (ICDL)* (pp. 1–2). San Diego: IEEE.
- Takahashi, D. Y., Narayanan, D., & Ghazanfar, A. A. (2013a). Development of self-monitoring essential for vocal interactions in marmoset monkeys. In *2013 IEEE 3rd Joint International Conference on Development and Learning and Epigenetic Robotics, ICDL 2013 - Electronic Conference Proceedings*. <https://doi.org/10.1109/DevLrn.2013.6652553>
- Takahashi, D. Y., Narayanan, D. Z., & Ghazanfar, A. A. (2013b). Coupled Oscillator Dynamics of Vocal Turn-Taking in Monkeys. *Current Biology*, 23, 2162–2168. <https://doi.org/10.1016/j.cub.2013.09.005>
- Teixidor, P., & Byrne, R. W. (1999). The “Whinny” of Spider Monkeys: Individual Recognition before Situational Meaning. *Behaviour*, 136(3), 279–308.

- Ter Maat, M., Truong, K. P., & Heylen, D. (2010). How turn-taking strategies influence users' impressions of an agent. In *Intelligent Virtual Agents: Proceedings of the 10th International Conference, IVA* (pp. 441–453). [https://doi.org/10.1007/978-3-642-15892-6\\_48](https://doi.org/10.1007/978-3-642-15892-6_48)
- Thierry, B. (2007). Unity in diversity: Lessons from macaque societies. *Evolutionary Anthropology*, 16(6), 224–238. <https://doi.org/10.1002/evan.20147>
- Tinbergen, N. (1963). On aims and methods of Ethology. *Zeitschrift für Tierpsychologie*, 20(4), 410–433.
- Toarmino, C. R., Wong, L., & Miller, C. T. (2017). Audience affects decision-making in a marmoset communication network. *Biology Letters*, 13(1), 20160934. <https://doi.org/10.1098/rsbl.2016.0934>
- Tobias, M. L., Viswanathan, S. S., & Kelley, D. B. (1998). Rapping, a female receptive call, initiates male-female duets in the South African clawed frog (*Xenopus laevis*; courtship songs; fertility advertisement; vocal behaviors; sexual selection). *Proceedings of the National Academy of Sciences*, 95, 1870–1875. [www.pnas.org](http://www.pnas.org).
- Tokuyama, N., & Furuichi, T. (2016). Do friends help each other? Patterns of female coalition formation in wild bonobos at Wamba. *Animal Behaviour*, 119, 27–35. <https://doi.org/10.1016/j.anbehav.2016.06.021>
- Tomasello, M. (2008). *Origins of Human Communication*. Cambridge: Massachusetts: MIT Press.
- Tomasello, M., Call, J., Nagell, K., Olguin, R., & Carpenter, M. (1994). The Learning and Use of Gestural Signals by Young Chimpanzees: A Trans-generational Study. *Primates*, 35(2), 137–154.
- Tomasello, M., George, B. L., Kruger, A. C., Jeffrey, M., Farrar, & Evans, A. (1985). The development of gestural communication in young chimpanzees. *Journal of Human Evolution*, 14, 175–186.
- Townsend, S. W., Deschner, T., & Zuberbühler, K. (2008). Female chimpanzees use copulation calls flexibly to prevent social competition. *PLoS ONE*, 3(6), e2431. <https://doi.org/10.1371/journal.pone.0002431>
- Townsend, S. W., Koski, S. E., Byrne, R. W., Slocombe, K. E., Bickel, B., Boeckle, M., et al. (2017). Exorcising Grice's ghost: an empirical approach to studying intentional communication in animals. *Biological Reviews*, 92(3), 1427–1433. <https://doi.org/10.1111/brv.12289>
- Toyoda, A., Maruhashi, T., Malaivijitnond, S., & Koda, H. (2017). Speech-like orofacial oscillations in stump-tailed macaque (*Macaca arctoides*) facial and vocal signals. *American Journal of Physical Anthropology*, 164(2), 435–439. <https://doi.org/10.1002/ajpa.23276>
- Trainer, J. M., McDonald, D. B., & Learn, W. A. (2002). The development of coordinated singing in cooperatively displaying long-tailed manakins. *Behavioral Ecology*, 13(1), 65–69. <https://academic.oup.com/beheco/article/13/1/65/181303>
- Tutin, C. E. (1996). Ranging and social structure of lowland gorillas in the Lope Reserve, Gabon. In & T. N. W. C. McGrew, L. F. Marchant (Ed.), *Great ape societies* (pp. 58–70). Cambridge: Cambridge University Press.
- Tutin, C. E. G., & Fernandez, M. (1993). Composition of the Diet of Chimpanzees and Comparisons With That of Sympatric Lowland Gorillas in the Lope Reserve, Gabon. *American Journal of Primatology*, 30(3), 30195–211.
- Tutin, C. E. G., Fernandez, M., Rogers, M. E., Williamson, E. A., & McGrew, W. C. (1991). Foraging profiles of sympatric lowland gorillas and chimpanzees in the Lope Reserve, Gabon. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 334(1270), 179–186.

- Upham, N. S., Esselstyn, J. A., & Jetz, W. (2019). Inferring the mammal tree: Species-level sets of phylogenies for questions in ecology, evolution, and conservation. *PLoS Biology*, *17*(12). <https://doi.org/10.1371/journal.pbio.3000494>
- Utami, S. S., & van Hooff, J. A. R. A. M. (1997). Meat-eating behavior of adult female orangutans. *American Journal of Primatology*, *43*, 159–165.
- van Elsacker, L., Vervaecke, H., & Verheyen, R. F. (1995). *A Review of Terminology on Aggregation Patterns in Bonobos (Pan paniscus)*. *International Journal of Primatology* (Vol. 16).
- van Hooff, J. A. R. A. M., & van Schaik, C. P. (1994). Male Bonds: Affiliative Relationships Among NonHuman Primate Males. *Behaviour*, 309–337.
- van Lawick-Goodall, J. (1968). The Behaviour of Free-living Chimpanzees in the Gombe Stream Reserve. *Animal Behaviour Monographs*, *1*(3), 165–311. [https://doi.org/10.1016/S0066-1856\(68\)80003-2](https://doi.org/10.1016/S0066-1856(68)80003-2)
- van Noordwijk, M. A., Arora, N., Willems, E. P., Dunkel, L. P., Amda, R. N., Mardianah, N., et al. (2012). Female philopatry and its social benefits among Bornean orangutans. *Behavioral Ecology and Sociobiology*, *66*, 823–834. <https://doi.org/10.1007/s00265-012-1330-7>
- van Schaik, C. P. (1999). The Socioecology of Fission-fusion Sociality in Orangutans. *Primates*, *40*(1), 69–86.
- van Schaik, C. P., Marshall, A. J., & Wich, S. A. (2009). Geographic variation in orangutan behavior and biology. Its functional interpretation and its mechanistic basis. In C. P. Wich, S.A., Utami-Atmoko, S.S., Setia, T.M., and van Schaik (Ed.), *Orangutans: Geographic variation in behavioral ecology and conservation* (pp. 351–361). New York: Oxford University Press.
- van Schaik, C. P., Preuschoft, S., & Watts, D. P. (2004). Great ape social systems. In D. R. Russon, A.E. and Begun (Ed.), *The evolution of thought: evolutionary origins of great ape intelligence* (pp. 190–209). Cambridge, UK: Cambridge University Press.
- Vauclair, J., & Cochet, H. (2016). Gestural communication: A royal road to language development. *Enfance*, *4*(4), 419–433.
- Vervaecke, H., De Vries, H., & van Elsacker, L. (1999). An Experimental Evaluation of the Consistency of Competitive Ability and Agonistic Dominance in Different Social Contexts in Captive Bonobos. *Behaviour*, *136*(4), 423–442.
- Vigilant, L., Hofreiter, M., Siedel, H., & Boesch, C. (2001). Paternity and relatedness in wild chimpanzee communities. *Proceedings of the National Academy of Sciences*, *98*(23), 12890–12895. [www.pnas.org/doi/10.1073/pnas.231320498](http://www.pnas.org/doi/10.1073/pnas.231320498)
- Vigliocco, G., Perniss, P., & Vinson, D. (2014). Language as a multimodal phenomenon: Implications for language learning, processing and evolution. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences*, *369*(1651), 20130292-undefined. <https://doi.org/10.1098/rstb.2013.0292>
- Voigt, C., Leitner, S., & Gahr, M. (2006). *Repertoire and Structure of Duet and Solo Songs in Cooperatively Breeding White-Browed Sparrow Weavers* (Vol. 143).
- von Helversen, D., Schul, J., & Kleindienst, H. U. (2001). Male recognition mechanism for female responses implies a dilemma for their localisation in a phaneropterine bushcricket. *Journal of Comparative Physiology - A Sensory, Neural, and Behavioral Physiology*, *186*(12), 1153–1158. <https://doi.org/10.1007/s003590000167>
- Waciewicz, S., & Żywicznyński, P. (2015). Language Evolution: Why Hockett’s Design Features are a Non-Starter. *Biosemiotics*, *8*(1), 29–46. <https://doi.org/10.1007/s12304-014-9203-2>
- Wadewitz, P., Hammerschmidt, K., Battaglia, D., Witt, A., Wolf, F., & Fischer, J. (2015).

- Characterizing vocal repertoires - Hard vs. Soft classification approaches. *PLoS ONE*, 10(4).  
<https://doi.org/10.1371/journal.pone.0125785>
- Waite, C., & Buchanan-Smith, H. M. (2001). What time is feeding? How delays and anticipation of feeding schedules affect stump-tailed macaque behavior. *Applied Animal Behaviour Science*, 75, 75–85.
- Wakefield, M. (2002). Female chimpanzee allogrooming behavior at Ngogo, Kibale National Park, Uganda. *American Journal of Physical Anthropology (Supplement)*, 34, 160–161.
- Warren, K. S., Verschoor, E. J., Langenhuijzen, S., Swan, R. A., Vigilant, L., & Heeney, J. L. (2001). Speciation and Intraspecific Variation of Bornean Orangutans, *Pongo pygmaeus pygmaeus*. *Mol. Biol. Evol.*, 18(4), 472–480. <https://academic.oup.com/mbe/article-abstract/18/4/472/979725>
- Waser, P. M., & Waser, M. S. (1977). Experimental Studies of Primate Vocalization: Specializations for Long-distance Propagation. *Zeitschrift für Tierpsychologie*, 43, 239–263.
- Watson, S. K., Townsend, S. W., Schel, A. M., Wilke, C., Wallace, E. K., Cheng, L., et al. (2015). Vocal learning in the functionally referential food grunts of chimpanzees. *Current Biology*, 25, 495–499. <https://doi.org/10.1016/j.cub.2014.12.032>
- Watts, D. P. (1996). Comparative socio-ecology of gorillas. In & T. N. W. C. McGrew, L. F. Marchant (Ed.), *Great Ape Societies* (pp. 16–28). Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/cbo9780511752414.004>
- Watts, D. P. (2002). Gorilla social relationships: A comparative overview. In *Gorilla biology: A multidisciplinary perspective* (pp. 302–327). Cambridge, United Kingdom: Cambridge University Press.
- Watts, D. P., & Mitani, J. C. (2001). Boundary Patrols and Intergroup Encounters in Wild Chimpanzees. *Behaviour*, 138, 299–327.
- Wells, K. D. (1977). The Social Behaviour of Anuran Amphibians. *Animal Behaviour*, 25, 666–693.
- Wernicke, C. (1874). *Der aphasische Symptomencomplex: eine psychologische Studie auf anatomischer Basis*. Cohn & Weigert.
- White, F. J. (1998). Seasonality and Socioecology: The Importance of Variation in Fruit Abundance to Bonobo Sociality. *International Journal of Primatology*, 19(6), 1013–1027.
- Wich, S. A., & De Vries, H. (2006). Male monkeys remember which group members have given alarm calls. *Proceedings of the Royal Society B: Biological Sciences*, 273(1587), 735–740. <https://doi.org/10.1098/rspb.2005.3320>
- Wich, S. A., Geurts, M. L., Setia, T. M., & Utami-Atmoko, S. S. (2006). Influence of fruit availability on Sumatran orangutan sociality and reproduction. In C. Hohmann, G., Rob-bins, M.M., and Boesch (Ed.), *Feeding ecology in apes and other primates. Ecological, physical and behavioral aspects* (pp. 335–356). Cambridge, UK: Cambridge University Press.
- Wich, S. A., Krützen, M., Lameira, A. R., Nater, A., Arora, N., Bastian, M. L., et al. (2012). Call cultures in orang-utans? *PLoS ONE*, 7(5), e36180. <https://doi.org/10.1371/journal.pone.0036180>
- Wich, S. A., Meijaard, E., Marshall, A. J., Husson, S., Ancrenaz, M., Lacy, R. C., et al. (2008). Distribution and conservation status of the orang-utan (*Pongo* spp.) on Borneo and Sumatra: How many remain? *Oryx*, 42(3), 329–339. <https://doi.org/10.1017/S003060530800197X>
- Wich, S. A., Swartz, K. B., Hardus, M. E., Lameira, A. R., Stromberg, E., & Shumaker, R. W. (2009a). A case of spontaneous acquisition of a human sound by an orangutan. *Primates*, 50(1), 56–64. <https://doi.org/10.1007/s10329-008-0117-y>
- Wich, S. A., Utami-Atmoko, S. S., Mitra Setia, T., Rijksen, H. D., Schürmann, C., van Hooff, J. A. R.



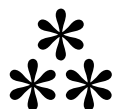
- A. M., & van Schaik, C. P. (2004). Life history of wild Sumatran orangutans (*Pongo abelii*). *Journal of Human Evolution*, 47(6), 385–398. <https://doi.org/10.1016/j.jhevol.2004.08.006>
- Wich, S. A., Utami-Atmoko, S. S., Setia, T. M., & van Schaik, C. P. (2009b). *Orangutans: Geographic variation in behavioral ecology and conservation*. New York: Oxford University Press.
- Wickler, W. (1980). Vocal Duetting and the Pair Bond I. Coyness and Partner Commitment. A Hypothesis. *Zeitschrift für Tierpsychologie*, 52, 201–209.
- Wiley, R. H. (2012). Communication as a transfer of information: measurement, mechanism and meaning. In U. E. Stegmann (Ed.), *Animal communication theory: information and influence* (pp. 113–129). Cambridge: Cambridge University Press.
- Wilke, C., Kavanagh, E., Donnellan, E., Waller, B. M., Machanda, Z. P., & Slocombe, K. E. (2017). Production of and responses to unimodal and multimodal signals in wild chimpanzees, *Pan troglodytes schweinfurthii*. *Animal Behaviour*, 123, 305–316. <https://doi.org/10.1016/j.anbehav.2016.10.024>
- Wilkinson, G. S. (2003). Social and vocal complexity in bats. In P. L. DeWaal, F. B. M. & Tyack (Ed.), *Animal social complexity: intelligence, culture, and individualized societies* (pp. 322–341). Cambridge: Harvard University Press.
- Williams, J. M., Hsien-Yang, L., & Pusey, A. E. (2002). Costs and benefits of grouping for female chimpanzees at Gombe. In G. H. L. F. M. C. Boesch (Ed.), *Behavioural diversity in chimpanzees and bonobos*. Cambridge: Cambridge University Press.
- Williamson, E. A., & Butynski, T. M. (2013). Gorilla gorilla Western Gorilla. In J. K. and J. K. T.M. Butynski (Ed.), *Mammals of Africa* (pp. 39–45). London, UK: Bloomsbury Publishing.
- Wilson, M. L., Boesch, C., Fruth, B., Furuichi, T., Gilby, I. C., Hashimoto, C., et al. (2014). Lethal aggression in *Pan* is better explained by adaptive strategies than human impacts. *Nature*, 513(7518), 414–417. <https://doi.org/10.1038/nature13727>
- Wilson, M. L., Hauser, M. D., & Wrangham, R. W. (2001). Does participation in intergroup conflict depend on numerical assessment, range location, or rank for wild chimpanzees? *Animal Behaviour*, 61(6), 1203–1216. <https://doi.org/10.1006/anbe.2000.1706>
- Winter, P., Handley, P., Ploog, D., Schott, D., Wintert, P., & Schott, D. (1973). Ontogeny of Squirrel Monkey Calls under Normal Conditions and under Acoustic Isolation. *Behaviour*, 47(4), 230–239.
- Wrangham, R. W., Clark, A. P., & Isabirye-Basuta, G. (1992). Female social relationships and social organisation of the Kibale Forest chimpanzees. In M. P. T. Nishida, W. C. McGrew, P. Marler & F. D. M. de Waal (Eds.), *Topics in primatology, Human origins (vol. 1, pp. 81–98)*. Tokyo: University of Tokyo Press.
- Wrangham, R. W., Gittleman, J. L., & Chapman, C. A. (1993). Constraints on group size in primates and carnivores: Population density and day-range as assays of exploitation competition. *Behavioral Ecology and Sociobiology*, 32, 199–209.
- Wrangham, R. W., & Smuts, B. B. (1980). Sex differences in the behavioural ecology of chimpanzees in the Gombe National Park, Tanzania. *Journal of Reproduction and fertility, Supplement*(Suppl 28), 13–31.
- Wright, T. F., & Dorin, M. (2001). Pair Duets in the Yellow-Naped Amazon (Psittaciformes: *Amazona auropalliata*): Responses to Playbacks of Different Dialects. *Ethology*, 107, 111–124.
- Xu, J., Gannon, P. J., Emmorey, K., Smith, J. F., & Braun, A. R. (2009). Symbolic gestures and spoken language are processed by a common neural system. *proceedings of the National Academy*

- of Sciences*, 106(49), 20664–20669. [www.pnas.org/doi/10.1073/pnas.0909197106](http://www.pnas.org/doi/10.1073/pnas.0909197106)
- Xue, Y., Prado-Martinez, J., Sudmant, P., Narasimhan, V., Ayub, Q., Szpak, M., et al. (2015). Mountain gorilla genomes reveal the impact of long-term population decline and inbreeding. *Science*, 348(6231), 242–245. <https://doi.org/10.1126/science.aaa4484>
- Yoshida, S., & Okanoya, K. (2005). Evolution of turntaking: a bio-cognitive perspective. *Cognitive Studies*, 12(3), 153–165.
- Yosida, S., Kobayasi, K. I., Ikebuchi, M., Ozaki, R., & Okanoya, K. (2007). Antiphonal vocalization of a subterranean rodent, the naked mole-rat (*Heterocephalus glaber*). *Ethology*, 113, 703–710. <https://doi.org/10.1111/j.1439-0310.2007.01371.x>
- Yosida, S., & Okanoya, K. (2009). Naked mole-rat is sensitive to social hierarchy encoded in antiphonal vocalization. *Ethology*, 115(9), 823–831. <https://doi.org/10.1111/j.1439-0310.2009.01677.x>
- Zimmermann, U., Rheinlaender, U., Irwin, J., & Robinson, D. (1989). Comparative Cues for male phonotaxis in the duetting bushcricket *Leptophyes punctatissima*. *Journal of Comparative Physiology A*, 164, 621–628.
- Zuberbühler, K., Ouattara, K., Bitty, A., Lemasson, A., & Noë, R. (2009). The primate roots of human language. In F. d’Errico & J.-M. Hombert (Eds.), *Becoming eloquent: Advances in the emergence of language, human cognition, and modern cultures* (Jhon Benjamins., pp. 235–264). Amsterdam. <http://www.springer.com/series/2506>
- Zuberbühler, K., & Wittig, R. M. (2011). Field experiments with non-human primates: A tutorial. In J. M. Setchell & D. J. Curtis (Eds.), *Field and laboratory methods in primatology: A practical guide* (pp. 207–224). Cambridge, U.K.: Cambridge University Press.
- Zumpe, D., & Michael, R. P. (1986). Dominance index: A simple measure of relative dominance status in primates. *American Journal of Primatology*, 10, 291–300. <https://doi.org/10.1002/ajp.1350100402>



# Chapitre VIII

## **Annexes**





## Annexe 1 | Liste des publications

### Chapitre d'ouvrage scientifique

**Pougnault, L.**, Levréro, F., Lemasson, A. (2020). Conversation among primate species. In N. Masataka (Ed.), *The Origins of Language Revisited - Differentiation from Music and the Emergence of Neurodiversity and Autism* (Pp. 73-96). Singapor: Springer.  
DOI : 10.1007/978-981-15-4250-3\_4

### Publication dans une revue internationale avec comité de lecture

**Pougnault, L.**, Levréro, F., Mulot, B., Lemasson A. (2020). Breaking conversational rules matters to captive gorillas: A playback experiment. *Scientific Reports* **10**, 6947.  
DOI :10.1038/s41598-020-63923-7

### Publications en préparation

**Pougnault, L.**, Lemasson A., Mulot, B., Levréro, F. Preliminary investigation of temporal calling patterns of a captive group of chimpanzees (en préparation)

**Pougnault, L.**, Levréro, F., Leroux, M., Paulet, J., Bombani, P., Dentressangle, F., Deruti, L., Mulot, B., Lemasson, A. «Conversational rules » are driven by social pressures, say studies on great apes (en préparation)



## Annexe 2 | Communications orales

### Communication orale invitée dans un colloque international

(Voir résumé page suivante)

**Pougnault L.**, Lemasson A., Mulot B., Levréro F. (2020) Social and temporal basis of vocal interactions in Western lowland gorillas. WorkShop EvoLang XIII, April 14-17, 2020, Brussels, Belgium (*reportée pour cause de crise sanitaire*)

### Communications orales dans des colloques nationaux

(Voir résumés page suivante)

**Pougnault L.**, Levréro F., Mulot B., Lemasson A. (2019). Importance du respect des « règles de conversations » chez les gorilles des plaines de l'Ouest (*Gorilla g. gorilla*). 32ème colloque de la Société Francophone de Primatologie (SFDP), 2-5 Octobre 2019, Parc de Branféré (56), France.

**Pougnault L.**, Levréro F., Mulot B., Lemasson A. (2020). Breaking conversational rules matters to captive gorillas: A playback experiment. Journées scientifiques de l'Ecole Doctorale Biologie Santé, 10-11 Décembre 2020, Rennes (35), France.

### Contribution à des séminaires de formations scientifiques

**Pougnault L.**, Levréro F., Mulot B., Lemasson A. (2019). The relevance of conversation rules in gorillas: a playback experiment. 2nd Edition Bioacoustic Winter School 2019, 7-18 Janvier 2019, Université de Lyon/Saint Etienne (42), France.

**Pougnault L.**, Lemasson A., Mulot B., Levréro F. (2020). Breaking conversational rules matters to captive gorillas? A playback experiment. 3ème Edition Bioacoustic Winter School 2019, 6-17 Janvier 2020, Université de Lyon/Saint Etienne (42), France.



**WorkShop EvoLang XIII****Social and temporal basis of vocal interactions in Western lowland gorillas**

Pougnault L., Lemasson A., Mulot B., Levréro F.

Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000 Rennes, France ; Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES / CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France ; ZooParc de Beauval & Beauval Nature, 41110 Saint Aignan, France

The psychologists and sociolinguists agree that, even though notable cultural variations exist, comparable conversational basic phenomena appear in unrelated languages (Beach, 1990). The universality of conversations has been well illustrated by the French social psychologist Rodolphe Ghiglione in his so-called “contract of communication” that conversing interlocutors informally agree on (Ghiglione, 1986). This contract facilitates vocal coordination between interlocutors and promotes inter-comprehension. Ghiglione (1986) listed several principles, amongst which “context relevance” (i.e. evaluation of context as pertinent or not to initiate a conversation), “reciprocity” (i.e. evaluation of the partner as a valid interlocutor), and “contract-based temporal rules” (i.e. the respect of the turn-taking which refers to a reciprocal exchange of alternating, short, and flexible turns between two or more interlocutors -Sacks et al., 1974- and speech overlap avoidance). Turn-taking represents the social core of human interactions (Levinson, 2016) and can be found in all languages and human societies around the world (Stivers et al., 2009). The fact that this temporal rule is universal among human cultures question its biological basis and motivates the search of parallels with the communicative interactions of nonhuman animals. Communicative turn-taking has been found in a broad range of nonhuman species (e.g. elephants *Loxodonta africana*: Leong et al., 2003; sperm-whales *Physeter macrocephalus*: Schulz et al., 2008; common bottlenose dolphins *Tursiops truncatus*: Janik, 2000; starlings *Sturnus vulgaris*: Henry et al., 2015; naked mole-rats *Heterocephalus glaber*: Yoshida & Okanoya, 2005; bats *Diaemus youngi*: Carter et al., 2008), including a number of nonhuman primates species (e.g. Campbell’s monkey *Cercopithecus campbelli*: Lemasson et al., 2011; common marmoset *Callitrix jacchus*: Takahashi et al., 2012; Japanese macaques *Macaca fuscata*: Sugiura & Masataka, 1995). However, few studies have focused on vocal turn-taking in nonhuman great apes and their results are contradictory (bonobos *Pan paniscus* Levréro et al., 2019; chimpanzee *Pan troglodytes*: Arcadi, 2000, but see Fröhlich et al., 2016 for gestural turn-taking in chimpanzee). So, the current state of knowledge prevents to conclude between a convergent evolution guided by the requirements of social life or a shared inheritance, suggesting an ancient mechanism which was already present. Authors also questioned the fact that these abilities to take turns in nonhuman animals are genetically programmed or, as in humans (Ginsburg & Kilbourne, 1988; Stern et al., 1975), socially acquired. Empirical evidence is currently lacking to answer this question, although the few available studies on monkeys showed that juveniles break conversational rules more often than do adults (howler monkeys *Alouatta pigra*: Briseño Jaramillo et al., 2017; Campbell’s monkeys: Lemasson et al., 2011). Nevertheless, playback experiments using a violation-of-expectation paradigm showed that adult monkeys audience clearly discriminated between appropriate and inappropriate vocal exchange patterns, whereas socially inexperienced juveniles did not (Campbell’s monkeys: Lemasson et al., 2011). This experiment suggests that the temporal organization of vocal interactions is not a simple neurobiological guided behaviour but a social awareness of the respect of rules.

Here, we propose to complete the knowledge about existence of turn-taking in great apes by looking at western lowland gorilla (*Gorilla gorilla gorilla*) through two complementary studies conducted on the same captive group. The first one, run by Lemasson and colleagues (2018), was to determine the social basis of vocal interactions on twelve individuals forming a unimale/multifemale group, with a social structure comparable with one found in the wild (Parnell, 2002). This first study found that individuals do seem to engage in socially- and temporally-ruled vocal interactions involving closed-calls named grunts (Hedwig et al., 2015; Hedwig et al., 2014; Salmi et al., 2014). During grunt exchanges, temporal pattern seems to be respected with an average “response” delay of 0.5 seconds (maximum 3sec) and with obvious call overlap

avoidance. Also preferred interlocutors are non-randomly chosen. Grunt exchanges are mainly dyadic interactions that are most frequently observed among partners close in age, regardless of kinship. To assess the social relevance of this two potential interactional rules (inter-call durations and the age difference of interlocutors) for the audience, we ran a violation-of-expectation paradigm using playback experiment (Pougnault al., under review). Attentional response showed us that call overlap avoidance matters to subjects, but the importance of the interlocutors' age proximity appeared less clear. The intensity of the response varied with subjects' age in a context-dependent way, supporting a possible role of learning. Our findings support the growing number of studies highlighting the importance of vocal turn-taking in animals and a possible sociogenesis of this ability in primates.

- Arcadi, A. C. (2000). Vocal responsiveness in male wild chimpanzees: Implications for the evolution of language. *Journal of Human Evolution*, 39(2), 205–223. <https://doi.org/10.1006/jhev.2000.0415>
- Beach, W. A. (1990). Searching for universal features of conversation. *Research on Language and Social Interaction*, 24(1–4), 351–368. <https://doi.org/10.1080/08351819009389347>
- Briseño Jaramillo, M., Guyot-Biquand, V., Estrada, A., & Lemasson Vocal, A. (2017). Vocal repertoire of free-ranging black howler monkeys' (*Alouatta pigra*): call types, contexts and sex-related contributions. *American Journal of Primatology*, 79(5), 22630–10. <https://doi.org/10.1002/ajp.22630>
- Carter, G. G., Skowronski, M. D., Faure, P. A., & Fenton, B. (2008). Antiphonal calling allows individual discrimination in white-winged vampire bats. *Animal Behaviour*, 76, 1343–1355. <https://doi.org/10.1016/j.anbehav.2008.04.023>
- Fröhlich, M., Kuchenbuch, P., Müller, G., Fruth, B., Furuichi, T., Wittig, R. M., & Pika, S. (2016). Unpeeling the layers of language: Bonobos and chimpanzees engage in cooperative turn-taking sequences. *Scientific Reports*, 6(25887), 1–14. <https://doi.org/10.1038/srep25887>
- Ghigliione, R. (1986). *L'homme communicant*. (A. Colin, Ed.). Paris, France.
- Ginsburg, G. P., & Kilbourne, B. K. (1988). Emergence of vocal alternation in mother-infant interchanges. *Journal of Child Language*, 15, 221–235. <https://doi.org/10.1017/S0305000900012344>
- Hedwig, D., Mundry, R., Robbins, M. M., & Boesch, C. (2015). Contextual correlates of syntactic variation in mountain and western gorilla close-distance vocalizations: Indications for lexical or phonological syntax? *Animal Cognition*, 18, 423–435. <https://doi.org/10.1007/s10071-014-0812-6>
- Hedwig, D., Robbins, M. M., Mundry, R., Hammerschmidt, K., & Boesch, C. (2014). Acoustic structure and variation in mountain and western gorilla close calls: A syntactic approach. *Behaviour*, 151, 1091–1120. <https://doi.org/10.1163/1568539X-00003175>
- Henry, L., Craig, A. J. F. K., Lemasson, A., & Hausberger, M. (2015). Social coordination in animal vocal interactions. Is there any evidence of turn-taking? The starling as an animal model. *Frontiers in Psychology*, 6, 1416. <https://doi.org/10.3389/fpsyg.2015.01924>
- Janik, V. M. (2000). Whistle Matching in Wild Bottlenose Dolphins (*Tursiops truncatus*). *Science*, 289, 1355–1357. <https://doi.org/10.1126/science.289.5483.1355>
- Lemasson, A., Glas, L., Barbu, S., Lacroix, A., Guilloux, M., Remeuf, K., & Koda, H. (2011). Youngsters do not pay attention to conversational rules: Is this so for nonhuman primates? *Scientific Reports*, 1. <https://doi.org/10.1038/srep00022>
- Lemasson, A., Pereira, H., & Levréro, F. (2018). Social Basis of Vocal Interactions in Western Lowland Gorillas (*Gorilla g. gorilla*). *Journal of Comparative Psychology*, 132(2), 141–151. <https://doi.org/10.1037/com0000105>
- Leong, K. M., Ortolani, A., Burks, K. D., Melleni, J. D., & Savage, A. (2003). Quantifying Acoustic and Temporal Characteristics of Vocalizations for a Group of Captive African Elephants *Loxodonta Africana*. *Bioacoustics The International Journal of Animal Sound and Its Recording*, 13, 213–231. <https://doi.org/10.1080/09524622.2003.9753499>
- Levinson, S. C. (2016). Turn-taking in Human Communication – Origins and Implications for Language Processing. *Trends in Cognitive Sciences*, 20(1), 6–14. <https://doi.org/10.1016/j.tics.2015.10.010>
- Levréro, F., Touitou, S., Frédet, J., Nairaud, B., Guéry, J.-P., & Lemasson, A. (2019). Social bonding drives vocal exchanges in Bonobos. *Scientific Reports*, 9, 711. <https://doi.org/10.1038/s41598-018-36024-9>
- Parnell, R. J. (2002). Group size and structure in western lowland gorillas (*Gorilla gorilla gorilla*) at Mbeli Bai, Republic of Congo. *American Journal of Primatology*, 56(4), 193–206. <https://doi.org/10.1002/ajp.1074>
- Pougnault, L., Levréro, F., Mulot, B., & Lemasson, A. (under review). Breaking conversational rules matters to captive gorillas: a playback experiment.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, 50, 696–735. <https://doi.org/10.2307/412243>
- Salmi, R., Hammerschmidt, K., & Doran-Sheehy, D. M. (2014). Individual distinctiveness in call types of wild western female gorillas. *PLoS ONE*, 9(7), e101940. <https://doi.org/10.1371/journal.pone.0101940>
- Schulz, T. M., Whitehead, H., Gero, S., & Rendell, L. (2008). Overlapping and matching of codas in vocal interactions between sperm whales: insights into communication function. *Animal Behaviour*, 76, 1977–1988. <https://doi.org/10.1016/j.anbehav.2008.07.032>
- Stern, D. N., Jaffe, J., Beebe, B., & Bennett, S. L. (1975). Vocalizing in unison and in alternation: Two modes of communication within the mother- infant dyad. *Annals of the New York Academy of Sciences*, 263(1), 89–100. <https://doi.org/10.1111/j.1749-6632.1975.tb41574.x>
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J., Yoon, K., & Levinson, S. C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106(26), 10587–10592. <https://doi.org/10.1073/pnas.0903616106>
- Sugiura, H., & Masataka, N. (1995). Temporal and acoustic flexibility in vocal exchanges of coo calls in Japanese macaques (*Macaca fuscata*). In E. Zimmermann, J. D. Newman, & U. Jürgens (Eds.), *Current topics in primate vocal communication* (NY : Plenum Press, pp. 121–140). New York. <https://doi.org/10.1007/978-1-4757-9930-9>
- Takahashi, D. Y., Narayanan, D., & Ghazanfar, A. A. (2012). A computational model for vocal exchange dynamics and their development in marmoset monkeys. In *International Conference on Development and Learning and Epigenetic Robotics (ICDL)* (pp. 1–2). San Diego: IEEE.
- Yoshida, S., & Okanoya, K. (2005). Evolution of turntaking: a bio-cognitive perspective. *Cognitive Studies*, 12(3), 153–165.

### **32ème colloque de la Société Francophone de Primatologie (SFDP)**

#### **Importance du respect des “règles de conversations” chez les gorilles des plaines de l’Ouest**

Pougnault L., Lemasson A., Mulot B., Levréro F.

*Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000 Rennes, France ; Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES / CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France ; ZooParc de Beauval & Beauval Nature, 41110 Saint Aignan, France*

Dans toutes les sociétés humaines, les conversations se structurent temporellement et socialement. L’universalité de ces règles de conversation et leurs apparitions précoces au cours de l’ontogénèse, suggèrent de possibles bases biologiques encourageant les comparaisons avec les interactions vocales décrites chez d’autres espèces animales. Ainsi, les interactions vocales temporellement et socialement organisées, qualifiées de proto-conversations, ont fait l’objet de nombreuses recherches sur les singes non-hominoïdes, mais contre toute attente, très peu d’études se sont intéressées aux grands singes. Le manque de connaissances à ce sujet empêche les chercheurs de dresser des scénari évolutifs solides concernant les potentielles origines des règles temporelles et sociales qui régissent nos conversations. Une étude récemment menée sur les gorilles de plaine de l’Ouest observe une organisation temporelle et sociale régissant les échanges de cris chez cette espèce, rappelant les règles de conversation humaine. En effet, les interactions vocales semblent majoritaires entre individus d’âge proche et les interlocuteurs s’avèrent respecter un délai inter-cris court, pouvant prévenir la superposition de leurs vocalisations. Au cours de notre propre recherche, nous avons testé la pertinence de ces potentielles règles pour six individus (trois femelles adultes et trois sub-adultes dont un mâle et deux femelles) évoluant au sein d’un groupe de gorilles captifs par la mise en place d’un paradigme expérimental de violation des attentes. Le principe de ce paradigme est de comparer la réaction spontanée d’un ou plusieurs individus à la présentation d’une situation habituelle en opposition à une situation inattendue, tel que le respect ou non des hypothétiques règles de conversation citées précédemment. Dans notre cas, l’objectif était de tester la pertinence sociale des potentielles règles conversationnelles pour l’audience. Pour ce faire, nous avons comparé la réaction spontanée des gorilles à l’encontre d’échanges de cris simulés de leurs congénères qui respectaient ou non le délai inter-cris court et la proximité d’âge des interlocuteurs. Les comportements d’attention visuelle, tels que l’orientation de la tête ou la durée du premier regard vers le stimulus, sont des indices essentiels permettant d’appréhender l’intérêt du sujet testé vis-à-vis de la stimulation qui lui est présentée. Suite aux diffusions, nous avons relevé que les individus portaient une attention accrue aux échanges vocaux respectant le délai inter-cris, ignorant ceux pour lesquels cette règle temporelle était bafouée, mais l’importance que portaient nos sujets à la proximité d’âge des interlocuteurs impliqués dans les échanges simulés était moins claire. Par ailleurs, l’intensité de l’attention que portaient les sujets aux stimuli variait en fonction de leur âge, ce qui peut sous-entendre un éventuel rôle de l’apprentissage sociale dans la perception de ces règles conversationnelles. Notre étude apporte une preuve supplémentaire de la pertinence sociale du respect de l’organisation temporelle et sociale des interactions vocales chez les animaux non-humains et de la possible sociogénèse de ces compétences chez les primates. La capacité à interagir vocalement a pu être une étape clé dans la co-évolution de la communication et d’une socialité complexe

Mots clés : interactions vocales, délai inter-cris, coupure de parole, proximité d’âge, grands singes

**Journées scientifiques de l'École Doctorale Biologie Santé**

**Breaking conversational rules matters to captive gorillas: A playback experiment.**

---

Pougnault L., Lemasson A., Mulot B., Levréro F.

*Univ Rennes, Normandie Univ, CNRS, EthoS (Éthologie animale et humaine) - UMR 6552, F-35000 Rennes, France ; Université de Lyon/Saint-Etienne, CNRS, Equipe Neuro-Ethologie Sensorielle, ENES / CRNL, UMR5292, INSERM UMR\_S 1028, Saint-Etienne, France ; ZooParc de Beauval & Beauval Nature, 41110 Saint Aignan, France*

---

Across human cultures, conversations are regulated by temporal and social rules. The universality of conversational rules suggests possible biological bases and encourages comparisons with the communicative interactions of nonhuman animals. Unexpectedly, few studies have focused on other great apes despite evidence of proto-conversational rules in monkeys, thus preventing researchers from drawing conclusions on potential evolutionary origins of this behaviour. A previous study showed however that western lowland gorillas engage in soft call interactions that seem temporally- and socially-ruled. Indeed, interactions occurred mainly between individuals close in age who followed a preset response delay, thus preventing call overlap. Here, we experimentally investigated the presence of these rules in a captive gorilla group, using a violation-of-expectation paradigm. Head orientation responses suggest that the respect of response delay matters to subjects, but the importance of the interlocutors' age proximity appeared less clear. The intensity of the response varied with subjects' age in a context-dependent way, supporting a possible role of learning. Our findings support the growing number of studies highlighting the importance of vocal turn-taking in animals and a possible sociogenesis of this ability. The capacity to "converse" might have been a key step in the co-evolution of communication and complex sociality.

Keywords: Vocal interactions, Turn-taking, Call overlap, Age, Great apes



### **Annexe 3 | Activités de diffusion de la culture scientifique**

**Intervention dans une classe de 4ème du Collège St Aubin d'Aubigné (35) :** présentation de l'éthologie et du laboratoire EthoS | Mars 2019

**Intervention dans une classe du collège St Yves à Mordelles (35) :** Présentation de l'éthologie et des possibles cursus en biologie animale | Mai 2019

**Interview pour la radio C-lab en amont du festival *Pint of Science* :** « La communication chez les primates non-humains » | Mai 2019

**Conférence pour le festival *Pint of Science* (Rennes) « Le langage est-il le propre de l'Homme ? : Sur les traces de l'origine de nos conversations »** | Mai 2019

**Création d'un jeu vidéoludique sur ordinateur dans le cadre du Festival Des Sciences :** « Comment survivre dans la jungle en écoutant les conseils des mones de Campbells ? » | Octobre 2019

**Interview en ligne pour le site de la station biologique de Paimpont :** « Retour sur le 32<sup>ème</sup> colloque de la Société Francophone de Primatologie », <https://paimpont.univ-rennes1.fr/actualites/retour-sur-le-32eme-colloque-de-la-societe-francophone-de-primatologie> | Octobre 2019

**Correction d'une intervention en médiation animale dans le cadre d'une ferme pédagogique à Laillé (35) :** « Présentation du comportement des animaux de la ferme : Qu'est-ce que le comportement et la science qui l'étudie ? » | Juin 2019

**Interview pour le 2<sup>nd</sup> numéro du magazine Beauval Nature :** « Les vocalisations chez les Grands Singes » | Juin 2020



## **Annexe 4 | Encadrement de projets, Responsabilité collectives & Formations suivies**

### Encadrements de stagiaires

Septembre 2019 – Novembre 2019 | **Co-encadrement d'un étudiant de Licence 2** – Julien Paulet, Licence Science de la Vie (Université Lyon 1) | Organisation temporelle des productions vocales des chimpanzés (*Pan troglodytes schweinfurthii*) d'une communauté sauvage de Budongo (Ouganda)

Octobre 2019 – Juin 2020 | **Co-encadrement d'un étudiant en Master 2** – Thomas Ferrari, Master Comportement Animal et Humain (Université Rennes 1) | Fonctionnalité et Directionnalité des échanges vocaux proto-conversationnels chez les gorilles des plaines de l'Ouest (*Gorilla g. gorilla*)

### Responsabilité collective au sein de l'unité

Mars 2019 – Novembre 2020 | **Création et gestion du « Journal Club » du laboratoire EthoS (UMR 6552)**

### Formations suivies et validées

Mai 2018 : **MOOC « Analyse des données multidimensionnelles »** | dispensée par AgroCampus Ouest (5 heures)

Juillet 2018 : **Statistiques – Introduction aux modèles mixtes linéaires et non-linéaires** | dispensée par Anastats (21 heures)

Octobre 2018 : **Expérimentation Animale de Niveau 1** | dispensée par l'Ecole Vétérinaire Oniris – Nantes (25 heures)

Novembre 2018 et mai 2019 : **Introduction à l'éthique de la recherche, Ethique en expérimentation animale et humaine** | dispensée par le Pôle Doctoral de Rennes (5 heures)

Avril 2019 : **Expérimentation animale et humaine** | dispensée par le Pôle Doctoral de Rennes (1 journée)







**Titre :** Sur les traces de l'origine de nos conversations : étude comparative des interactions vocales chez les grands singes

**Mots clés :** Grands singes, conversations, échanges vocaux, tour de parole, systèmes sociaux

**Résumé :** L'universalité des "règles" qui régissent les conversations humaines, comme le respect du tour de parole, suggère de possibles bases biologiques à ce comportement. Au travers d'une revue de la littérature, nous avons révélé qu'à l'image de nos conversations, les échanges vocaux de vocalisations de contact de la plupart des primates non-humains respectent des règles sociales et temporelles. Notamment, les interlocuteurs, choisis sur des critères sociaux, alternent leurs émissions vocales tout en évitant de se couper la parole. Ces échanges permettent de créer et consolider des relations affiliatives. Toutefois, les grands singes sont peu représentés dans la littérature disponible sur ce sujet, malgré l'intérêt comparatif qu'ils représentent de par leur proximité phylogénétique avec l'humain et la diversité de leurs systèmes sociaux. Ce projet vise donc à décrire les règles d'interactions vocales chez les grands singes non-humains : bonobos (*Pan paniscus*), chimpanzés (*Pan troglodytes*), gorilles des plaines de l'Ouest (*Gorilla gorilla gorilla*) et orangs-outans de Bornéo (*Pongo pygmaeus*). Dans un premier temps, une approche expérimentale a permis de vérifier l'importance du respect du tour de parole lors des échanges vocaux pour les gorilles et de

souligner un rôle possible de l'apprentissage social dans l'acquisition des règles appropriées d'interaction. Ensuite, une approche observationnelle menée sur un groupe de chimpanzés captifs a, au contraire, souligné l'extrême rareté des tours de paroles, au profit d'interactions vocales de type *chorus* et de fréquentes émissions de cris isolés associées à la régulation des conflits. Enfin, une approche comparative des quatre espèces, mêlant des données provenant de la littérature à nos propres résultats, a suggéré que l'organisation temporelle des interactions vocales est sous l'influence de caractéristiques sociales telles que le degré de socialité et de tolérance intra-groupe, comme le souligne l'existence d'échanges vocaux organisés chez les gorilles et les bonobos. La découverte d'échanges vocaux organisés chez les grands singes permet de compléter l'arbre phylogénétique de cette capacité langagière. Des investigations comparatives supplémentaires restent à mener afin d'affiner nos connaissances quant aux multiples et subtils facteurs sociaux ayant favorisé l'émergence de ce comportement interactionnel.

**Title :** In search of the origin of our conversations: a comparative study of vocal interactions in great apes

**Keywords :** Great apes, conversations, vocal exchanges, turn-taking, social systems

**Abstract :** The universality of the "rules" that govern human conversations, such as the respect for turn taking, suggests possible biological bases for this behaviour. Through a review of the literature, we revealed that like our conversations, the vocal exchanges of contact calls of most non-human primates respect social and temporal rules. In particular, the interlocutors, chosen on the basis of social criteria, alternate their vocal productions while avoiding overlapping each other's calls. These exchanges allow for the creation and consolidation of affiliative relationships. However, great apes are poorly represented in the available literature on the topic, despite the comparative interest they represent due to their phylogenetic proximity to humans and the diversity of their social systems. This project aims to describe the rules of vocal interactions in non-human great apes: bonobos (*Pan paniscus*), chimpanzees (*Pan troglodytes*), Western lowland gorillas (*Gorilla gorilla gorilla*) and Bornean orang-utans (*Pongo pygmaeus*). First, an experimental approach allowed us to ascertain that the turn-taking principle matters to gorillas during

vocal exchanges, and to highlight a possible role of social learning in the acquisition of appropriate rules of interaction. Secondly, an observational approach conducted on a group of captive chimpanzees highlighted, on the contrary, the extreme rarity of turns of taking, to the benefit of *choruses* and frequent emissions of isolated calls associated with the regulation of conflicts. Finally, a comparative approach of the four species, mixing data from the literature with our own results, suggested that the temporal organisation of vocal interactions is under the influence of subtle social characteristics, such as the degree of sociality and intra-group tolerance, as underlined by the existence of organised vocal exchanges in gorillas and bonobos. The discovery of organised vocal exchanges in great apes allow us to complete the phylogenetical tree of this ability. Further comparative work remains to be done in order to refine our knowledge about the multiple and subtle social factors that have favoured the emergence of this interactional behaviour.