SINGING BEHAVIOUR OF WHITE-HANDED GIBBONS (Hylobates lar) OF KHAO YAI NATIONAL PARK, THAILAND

PhD thesis submitted to the faculty of Science
Institute of Biology
University of Neuchâtel

For the degree of PhD in Biology

By

Julie Andrieu

Accepted by the expert committee:

Prof. Klaus Zuberbühler, University of Neuchâtel, thesis director
Prof. Redouan Bshary, University of Neuchâtel, internal examiner
Prof. Simon Townsend, University of Zurich, external examiner

Defended on the 31st of January 2020

UNIVERSITÉ DE
NEUCHÂTEL

Faculté des Sciences
Secrétariat-décanat de Faculté
Rue Emile-Argand 11
2000 Neuchâtel – Suisse
Tél : + 41 (0)32 718 21 00
E-mail : secretariat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Madame Julie Andrieu

Titre:

"Singing behaviour of white-handed gibbons (*Hylobates lar*) of Khao Yai National Park, Thailand"

sur le rapport des membres du jury composé comme suit:

- Prof. Klaus Zuberbühler, directeur de thèse, Université de Neuchâtel, Suisse
- Prof. Redouan Bshary, Université de Neuchâtel, Suisse
- Prof. Simon Townsend, Université de Zürich, Suisse

Neuchâtel, le 21 août 2020

Le Doyen, Prof. A. Bangerter

GENERAL ABSTRACT

This thesis manuscript focusses on the remarkable vocal abilities of white-handed gibbons (*Hylobates lar*) of Khao Yai National Park, Thailand. To understand the mechanisms underlying the acoustic communication system of these gibbons, I conducted my research project in Khao Yai National Park in Thailand, on 13 wild groups. White-handed gibbons, also known as lar gibbons, part of the Hylobatidae family, are the evolutionarily most distant relatives of humans within the hominoid lineage and are of interest regarding their complex vocal communication system. Their large vocal repertoire, lengthy songs and abilities to produce different song types - resultant from multiple note assemblages - make them an exceptional study system to address scientific questions, also about the origin of human vocal communication. A first well-known feature is that the mated pair sings stereotyped, loud and long songs on a daily basis and in a synchronised fashion; the duet songs. These songs are thought to function primarily in pair-bonding and mate and territorial defence, but these hypotheses have never been investigated in much depth. Another well-known feature is that gibbon groups also sing in response to predators and from previous research it became apparent that predator songs are assembled in different ways compared to duet songs. Here also, not much has been done beyond these basic findings.

A first goal of this research was therefore to investigate the intergroup singing interactions and explore the role of duet songs as a complex territorial communication network. Duet songs often spread to neighbouring groups, resulting in song waves. In my research I have been able to show that duet responses can either non-aggressive turn-taking exchanges, where groups delay their song onset so that the preceding gibbon group can end their duet without interruption; alternatively, duet responses can be competitive, counter-singing exchanges, in which case the song of the preceding group is temporally overlapped by the subsequent duet song. To investigate this phenomenon, I examined the mechanisms underlying the duet song production of two subsequent groups. Results showed that intergroup distance, group composition and, most importantly, genetic relatedness had a significant impact on whether gibbons opted for competitive or non-aggressive duet exchanges with their neighbours. The study thus shows that kinship plays a role in non-aggressive interactions between groups.

The second part of my research focusses on gibbon songs produced in response to predators, particularly how the same basic song units are assembled in context-specific ways. I was interested in how gibbons use their elaborate songs to encode external events and what inferences listeners can make from these vocalizations. Previous observations have shown that neighbouring groups respond to songs with the matching type, i.e. neighbours' duets by their own duets, but also predator songs by their own predator songs. This happens in the absence of visual contact, suggesting that gibbons can decode and understand the meaning of each other's songs, similar to how humans use language to address and understand each other. To investigate these coding and inference abilities in gibbons, I conducted playback and predator model experiments. In a first study, I broadcasted predator and duet songs to different groups and recorded vocal and behavioural responses. In a second study, I presented two predator types, the clouded leopard and the reticulated python, and again recorded subjects' vocal and behavioural responses.

Regarding the playback experiments, I found that gibbons could discriminate and infer the meaning of different song types (i.e. duet song and predator song) by displaying different vocal and behavioural responses. Regarding predator model presentations, a previous study showed that the leopard model reliably elicited strong vocal and anti-predatory behavioural responses, unlike the python model. This time, the life-sized python model also elicited strong singing and anti-predator responses, in contrast to previous research. More importantly, gibbons arranged song elements differently and songs could be distinguished regarding several acoustic components and sequential arrangements according to the type of predator encountered, a basic requirement for referential alarm calling behaviour. The basal position of gibbons in the Hominoïdae family combined with their usually complex vocal abilities make white-handed gibbons a particularly promising species to address questions of how non-human primates code external events in their vocal signals and what inferences they can make from these vocalizations. This kind of data will be relevant to better understand the conditions that might have led to the emergence and current complexity of human vocal communication, as seen in spoken language.

Keywords: White-handed gibbons, Duet song, Intergroup interactions, Turn-taking, Counter-singing, Kinship, Predator song, Playback experiments, Predation, Phonology

RÉSUMÉ

Ce manuscrit de thèse est axé sur les capacités vocales singulières des gibbons à mains blanches (*Hylobates lar*) du parc national de Khao Yai, en Thaïlande. Afin d'appréhender les mécanismes sous-jacents de leur système de communication, ce projet de recherche a été mené dans le parc national de Khao Yai en Thaïlande, au sein d'une population de gibbons dans leur environnement naturel, sur 13 groupes habitués à la présence humaine. Les gibbons à mains blanches, aussi appelés gibbon lar, font partie de la famille des Hylobatidés. Ils sont phylogénétiquement, les parents les plus éloignés de l'Homme au sein de la super-famille des Hominoïdae. L'étude de leur système complexe de communication s'avère donc, d'un très grand intérêt. Ils possèdent un répertoire de notes conséquent, produisent des chants de longue durée et ont la capacité de produire différents types de chants - résultant de la combinaison de plusieurs notes – ces attributs faisant d'eux un sujet d'étude unique pour aider à répondre à des problématiques scientifiques, telle que les origines du langage humain.

Une caractéristique bien connue concernant les capacités vocales des gibbons, réside dans la production quotidienne et en synchronie, par le couple adulte mâle et femelle de chaque groupe, de chants stéréotypés longs et puissants ; qu'on appelle duos. On suppose que ces chants ont pour fonctions principales le renforcement du lien entre les individus du couple adulte, ainsi que la défense du territoire et du partenaire, cependant, ces hypothèses n'ont jamais été explorées de manière détaillée. Un autre aspect de leur production vocale concerne la production de chants prédateurs, qui, par le biais d'une précédente étude, présentent des différences structurelles notables qui les différencient des duos. Sur ce point également, nous en savons encore que très peu au-delà de ce postulat de base.

L'objectif premier de ce projet de recherche était d'étudier les échanges interactifs de duos entre groupes de gibbons et d'explorer le rôle de la production de ces duos en tant que réseau de communication territorial complexe. Les duos sont généralement contagieux, se propageant aux groupes voisins, entrainant des vagues de duos dans la population. Cette étude a permis de démontrer que les duos émis en réponse à un duo précédant pouvaient être échangés en alternance de manière non-agressive, c'est-à-dire que les groupes vont temporiser le début de leur propre chant de sorte que le groupe précédent puisse terminer son duo sans être interrompu. Ces réponses pouvaient aussi être échangées de manière compétitive, en contre-chant, et dans ce cas les groupes vont initier leur propre duo par-dessus le duo précédent. Afin d'étudier ce phénomène, j'ai analysé les mécanismes régissant les échanges

de duos entre deux groupes consécutifs. Les résultats obtenus ont montré que la distance entre les groupes, la composition sociale du groupe, et surtout le lien de parenté génétique entre les individus des deux groupes, avaient un effet significatif sur les modalités d'échange de duos entre groupes (contre-chant ou alternance). Cette étude a ainsi permis de mettre en évidence le rôle de la parenté dans les interactions vocales entre groupes, facilitant les échanges de manière non compétitive.

La deuxième partie de ce projet de recherche est axée sur la production de chants prédateurs, et plus spécifiquement, sur la manière qu'ont les gibbons d'assembler différemment les mêmes éléments acoustiques de base de manière contextuelle, d'un type de chant à un autre. Je me suis donc intéressée à la façon dont les gibbons utilisent leurs chants sophistiqués pour encoder des événements externes et aux inférences que les individus réceptionnant ce signal peuvent extraire de ces vocalisations. De précédentes observations ont relaté la capacité des groupes de gibbons à répondre par le type de chant correspondant, c'est-à-dire à produire un duo en réponse à un duo émis par un groupe voisins, et de la même façon, à répondre par un chant prédateur en réponse à un autre chant prédateur. Ce phénomène se produisant en l'absence de contact visuel, sous-tend que les gibbons peuvent décoder et donc déduire la signification des différents types de chants produits par d'autres groupes, similaire à la façon dont les humains utilisent le langage pour communiquer et se comprendre. Afin d'étudier leurs capacités à produire des chants spécifiques et leurs capacités d'inférence, j'ai mené des expériences de playback et de présentation de modèles de prédateurs. Dans la première étude, j'ai diffusé, via un haut-parleur, des chants prédateurs ou des duos à différents groupes, et collecté leurs réponses vocales et comportementales. Dans la deuxième étude, j'ai présenté à plusieurs groupes différents modèles prédateurs, la panthère nébuleuse et le python réticulé, et de la même façon, j'ai collecté leurs réponses vocales et comportementales.

Pour ce qui est des expériences de playback, les résultats ont démontré que les gibbons pouvaient discriminer et extraire des informations différentes de deux types de chants, duos et chants prédateurs, en manifestant différentes réponses vocales et comportementales. Concernant les présentations de modèles prédateurs, une étude précédente a montré que le modèle panthère suscitait chez les gibbons, et de manière fiable, de fortes réponses vocales et des comportements anti-prédateurs, contrairement au modèle python. Cette fois, le modèle de python a aussi provoqué de fortes réponses vocales et des comportements anti-prédateurs. Mais surtout, les analyses acoustiques ont montré que les divers types de chants produits

pouvaient être différenciés par plusieurs structures acoustiques et arrangements de séquences, en fonction du type du type de prédateur, un critère de base de production de signaux d'alarmes référentiels.

Les Hylobatidés, comprenant les gibbons, sont les premiers à avoir divergé de la famille des Hominoïdae. Leur position phylogénétique et leurs capacité vocales complexes, font des gibbons à mains blanches une espèce particulièrement prometteuse pour tenter de comprendre et d'expliquer comment les primates non-humains ont la capacité de coder leurs productions vocales en fonction du contexte, et les inférences qu'ils peuvent extraire de ces vocalisations. Ce projet a pour but d'apporter des données et des éléments pertinents, afin de mieux comprendre les conditions qui ont pu mener à l'émergence et à la complexité du langage humain.

Mots-clés : Gibbons à mains blanches, Duo, Interactions entre groupes, Tour de parole, Contre-chant, Parenté, Chant prédateur, Expériences de playback, Prédation, Phonologie

ACKNOWLEDGEMENTS

First, I would like to thank Klaus Zuberbühler for this great PhD opportunity. His supervision, precious comments and support had made this project finally comes to light. This experience would not have been either possible without the precious help and professional support of Prof. Ulrich. H. Reichard and Prof. Suchinda Malaivijitnond. They provided all the administrative, institutional and official assistance required to make the access to the study site and thus my data collection possible. I also would like to deeply thank, collaborator and friend, Dr ChRistof Neumann, who patiently taught me and helped me with statistics. By the way, it is significantly your fault if I am not able to perform any statistical analysis without using R now… A special thanks to my field assistant and also collaborator, Dr Sam Penny, for his valuable help while conducting experiments.

Back from field, I could meet wonderful colleagues, with whom I have shared unforgettable time, thank you Coco, Gladez and Stephanie and for a more masculine touch, Christophe and Pawel. And of course, a special mention to my "Twisted sisters" Claudia, Noémie, Emilie and Yaëlle.

Finally, I would like to thank my close friends: Karine, Lolo, JC and Romain; my unforgettable Master degree 'Ecophy' buddies: Marion, Aurélie, Chouchou, Céline et Stéphane; and especially my family, for their unwavering support and help during my PhD. One of my experiment would not even have been possible without the help of my mother and my step-granny, who did the most of work in the elaboration of "our" python model, that they manufacture twice! Because it was so resembling, that the first one was stuck at the airport custom in Thailand and only released when half of the experiments were achieved (with the second python). Thanks for making me the master of 'python' puppet! Again, my mother, Cathy, for her precious help and undeniable support, she was assuredly proud of me but also worried every second of my time spent in the forest (and out of my hometown). She could not refrain herself from sending me food even if this time, I was not starving! I would like to thank my father, Gilles, who genetically gave me the patience to overcome and relativize all problems, and who was every time worried about my progress and future. Thanks to my stepfather, Hervé and stepmother, Nicole, for their precious support, making me believe any time that I could achieve anything in life. A special thanks to my brothers, Jimi and Axel, because whatever they do or say, they make my day better and they always made me feel I

was a special sister. I also want to thank, my mother and father in law, for all their support and for giving me extra-time to finalise my manuscript, and by amongst other things, off-loading my family duties. Last, but not least, I want to thank my best friend, the best partner in the world, who appeased me during hard times, always brought support and reassured me during this whole process. He also provided me with objective, new and naive ideas about all the projects I finally achieved when I was stuck. Without him and our little sunshine, it would have been much more difficult that it finally was. Thank you so much Frédéric 'Baloo' Bel and Alexandre 'Pipou' Bel.

Even if it comes to an end, this amazing human journey, left incredible and unforgettable memories that I will certainly tell about all my life long… I encountered, and was so close of, many wild animal species that, it is more than a feeling, I don't stop believe in thinking that I was considered to be part of their habitat. My childhood dream was to be like these 'guys', when I was watching animals' documentaries, who were into the forest spotting as many animals as they could see. In the end, I did it, I was welcome to the jungle and at the time, nothing else mattered.

I dedicate this thesis to my grand-mother, Bernadette, who is 'my' most wonderful granny in the whole world. She devoted her life to raise her children and gran-children without having any possibilities to work, 'Mamie', THIS work is as much yours as it is mine.

Table of Contents

3. White-handed gibbons discriminate context-specific song compositions 59

4. Gibbon songs refer to external events 87

1- GENERAL INTRODUCTION

Investigating the origins of human language and the evolutionary processes responsible for the emergence of our sophisticated communication system is a fascinating research topic. The prerequisites for human language have probably at least partially evolved from a common ancestor, shared with our phylogenetically closest non-human primate relatives. For this reason, many studies have scrutinised the communication system of different primate species for the purpose of identifying the evolutionary processes that could have led to the complex communication skills seen in humans. Research has mainly targeted our closest relatives, the great apes, but Hylobatids are an equally an interesting model to investigate the roots of human language, due to their basal position in the hominoid lineage.

Throughout this manuscript, I will focus on the vocal abilities of one Hylobatid species, the white-handed gibbon (*Hylobates lar*). My research will focus on their singing behaviour, hereby covering questions about their flexibility during song exchanges with neighbours and their capacity of producing and discriminating contextually distinct song types.

1.1 White-handed gibbon song repertoire

History

The first documented appearance of the loud and conspicuous songs of gibbons date back the 4th century in the Chinese writings: "***Sad the calls of the gibbons at the gorges of Pa-tung; after three calls in the night tears wet the traveller's dress***" (Chinese song, 4th Century, cited in Van Gulik, 1967, p.46). The scientific study of gibbon communication skills did not start before the 1930s. Carpenter, (1940) was first to record the calls of gibbons, with a first attempt to identify and classify their songs. Since then, many other researchers focussed on the study of gibbon vocal production. In particular, Haimoff was first to review and gather all the information about the acoustic structures produced by all gibbons species and to standardise the terminology of gibbon songs (Haimoff, 1984a).

From the 1980s, contributions about vocal behaviour of white-handed gibbons (*Hylobates lar*) became frequent especially by Raemaekers and collaborators (1984a), who also used playback methods. Studies were made on the mysterious and highly contagious "Ooaa" duet, on the mechanisms and function of long-range duet songs, and on the production of 'disturbed calls' given to alarm calls of other species, potential predators, including human observers, or during intra-or intergroup conflicts (Raemaekers and Raemaekers, 1984b, 1985a, 1985b). Finally, Clarke and collaborators focussed on the vocal production of white-handed gibbons by presenting predator models, which revealed new patterns of song production and their associated anti-predatory behaviours (Clarke et al., 2006, 2012, 2015).

Terminology

To remain consistent with previous studies on white-handed gibbons, I will use the terminology and definitions of gibbon song used by Raemaekers and Raemaekers and other researchers (Haimoff, 1983; Raemaekers et al., 1984a; Geissmann, 1993). The main components of gibbon song are notes, figures, phrases and sequences.

A ***'Note'"*** is a continuous tonal sound, of any distinct frequency or frequency modulation, emitted by one individual, produced during either inhalation or exhalation.

A ***'Figure'*** is a distinctive assemblage of single notes uttered in series, or a combination of a single note and an exhaling event.

A ***'Phrase'*** comprises a 'larger and looser' succession of notes separated by short pauses, produced in an idiosyncratic manner by a single individual. Phrases, themselves, are delimited by longer pauses.

A ***'Song'*** is defined as a collection of all the notes produced, from the first to the last, by one individual (e.g. solo song), two (e.g. duet songs) or more (e.g. disturbed songs or predator songs) with, arbitrary defined, no period of silence of more than ten minutes between notes (Haimoff, 1983).

A ***'Duet song'*** is defined as a song bout in which individuals of the mated pair interactively combine their vocalisations so that their respective vocal contributions are mutually coordinated."

a. Notes

All white-handed gibbons notes range in frequency from 400 to 1,600 Hertz. To date, eight types of basic note have been identified and associated to singing contexts (Table 1.1 and Figure 1.1). Six of them were carefully described by Raemaekers and collaborators back in the 80s. Then, the 'hoo' note, originally included in another existing note class (i.e. in 'wa' notes), has been redefined as a class of its own. An eighth class of note, designated as '**other notes'**, not comparable to any other description seen before was identified by Clarke and collaborators (2006). 'Other notes' comprises two different note shapes that I classified as 'other note type I' and 'other note type II', their specific production context and function are to date unknown. All notes mostly start by increasing in frequency, containing no more than two inflections, except for 'other note type II". Male vocal contributions are typically higher in pitch than female contributions. All basic notes can additionally be subject to pitch modulations that are divided into three classes (Raemaekers et al., 1984a):

- *Tremble*: this modulation can be found in notes produced by both males and females and is thought to be imputed to a lack of voice control which cause is currently unknown. It sounds like shallow and irregular modulation in notes of long duration.

- *Quaver*: this smooth, deep and regular modulation is only performed by the male, and mainly during the production of '*waoo'* notes. Contrary to the tremble, the quaver is thought be vocally controlled.

- *Wooble*: this modulation is acoustically situated between trembles and quavers. It is defined as a "brief reversal of reduction in the rate of change of pitch" (Raemaekers et al., 1984a).

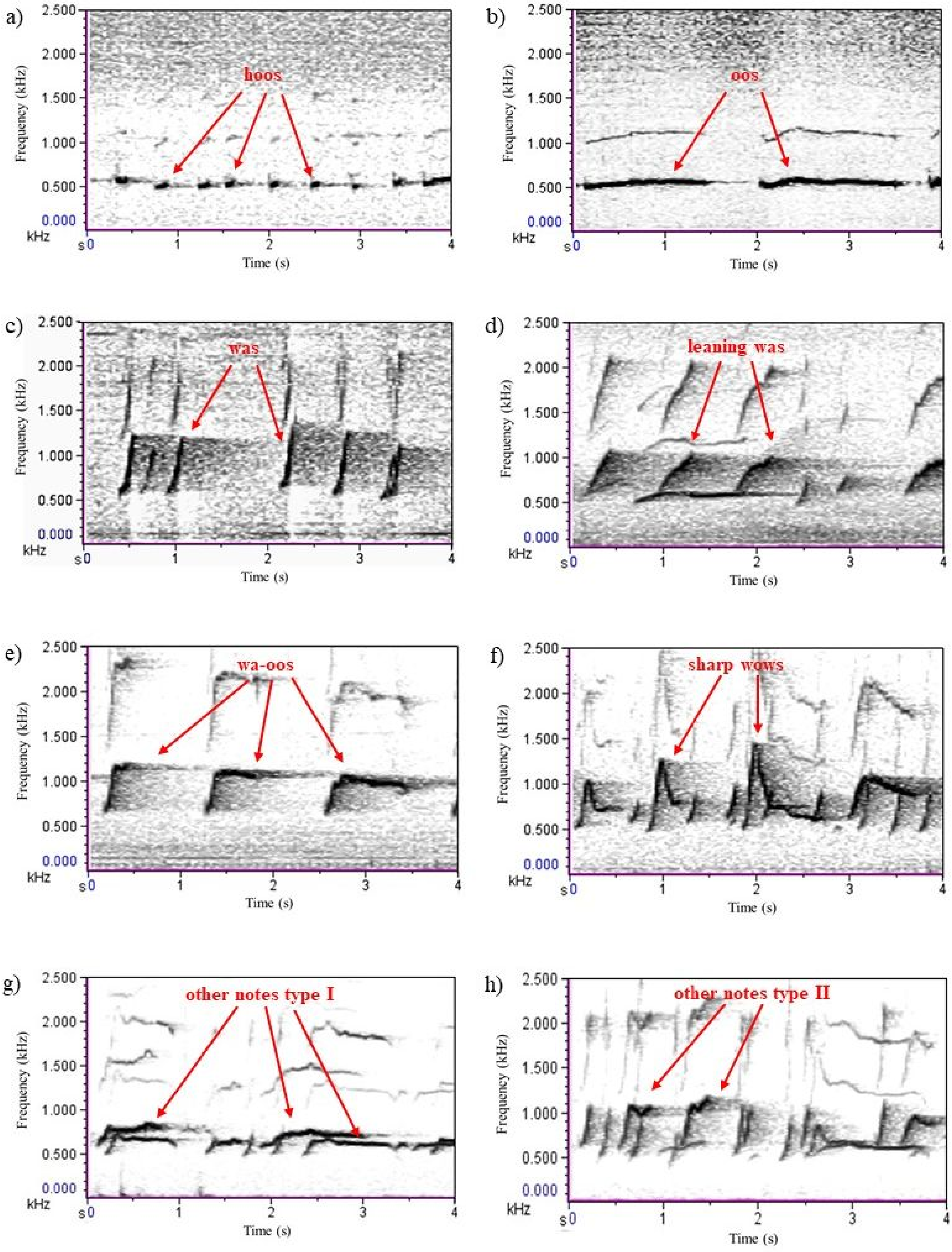

Figure 1.1. Song note repertoire of white-handed gibbons (Raemaekers et al., 1984a; Clarke et al., 2006;): a) ‘hoo’ note; b) ‘oo’; c) ‘wa’; d) ‘leaning wa’; e) ‘wa-oo’; f) ‘sharp wow’; g) ‘other note type I’ and h) ‘other note type II’. Songs were digitized using Cool Edit Pro 2.1. spectrograms were drawn using 21.6 Hz filter bandwidth, 2.69 Hz frequency resolution, 33.3 ms time grid resolution and a Hanning window function. ‘Ooaa’ notes are not illustrated here, they will be further described in section 1.1.b as part of the ‘Ooaa’ figure description and illustrated in fig 1.2b).

Table 1.1 The note repertoire of white-handed gibbons (Raemaekers et al., 1984a; Clarke et al., 2006; Andrieu, unpublished data)**.**

Note type	Description	Context
Hoo	Discrete and short note that is of low frequency and comprises within a narrow range of frequency. Produced by all group members	In introductory sequence of duet and disturbed call bouts
Oo	Note that is of a low, even frequency, rising slightly at first. It is mainly a female note	Mostly in the introductory sequence of a duet and in disturbed call bouts.
Wa	Short and uninflected note that sharply rises either in a straight or slightly concave way. Given by both, male and female	In all song bouts
Leaning wa	A note that resembles to the wa but longer in duration and leaning more to the right. Given by both, male and female	Introductory sequence of duet song
Wa-oo	Plaintive inflected note rising steeply, holding more or less the peak frequency to end up in a horizontal line of varying duration, it may decrease in frequency again as in a convex, dome-shaped note. Sung by either sex	Mostly in disturbed song bouts, in the introductory sequence of duet song bouts and in contact call bouts
Ooaa	Ooaa is a highly concave *wa* note, that ends up being J-shaped. It can be produced by all group members (see Figure 1.2.b)	In Ooaa duets mainly
Sharp wow	A note that is loud and penetrating that rises quickly to a high frequency, and then inflects sharply to fall in a concave curve. Mostly uttered by adult and sub-adult males	Mostly in disturbed song bouts, sometimes in the introductory sequence of duet song bouts in counter-singing exchanges
Others	Other note type I is a prolonged note of general low frequency subjected to tremble pitch modulation, slightly rising to form an initial bump-shape. Other note type II is a short note containing two inflections. Type II sharply rises, then inflects by decreasing slightly in frequency, and inflects again to higher frequencies.	In disturbed song bouts, but also in duets songs

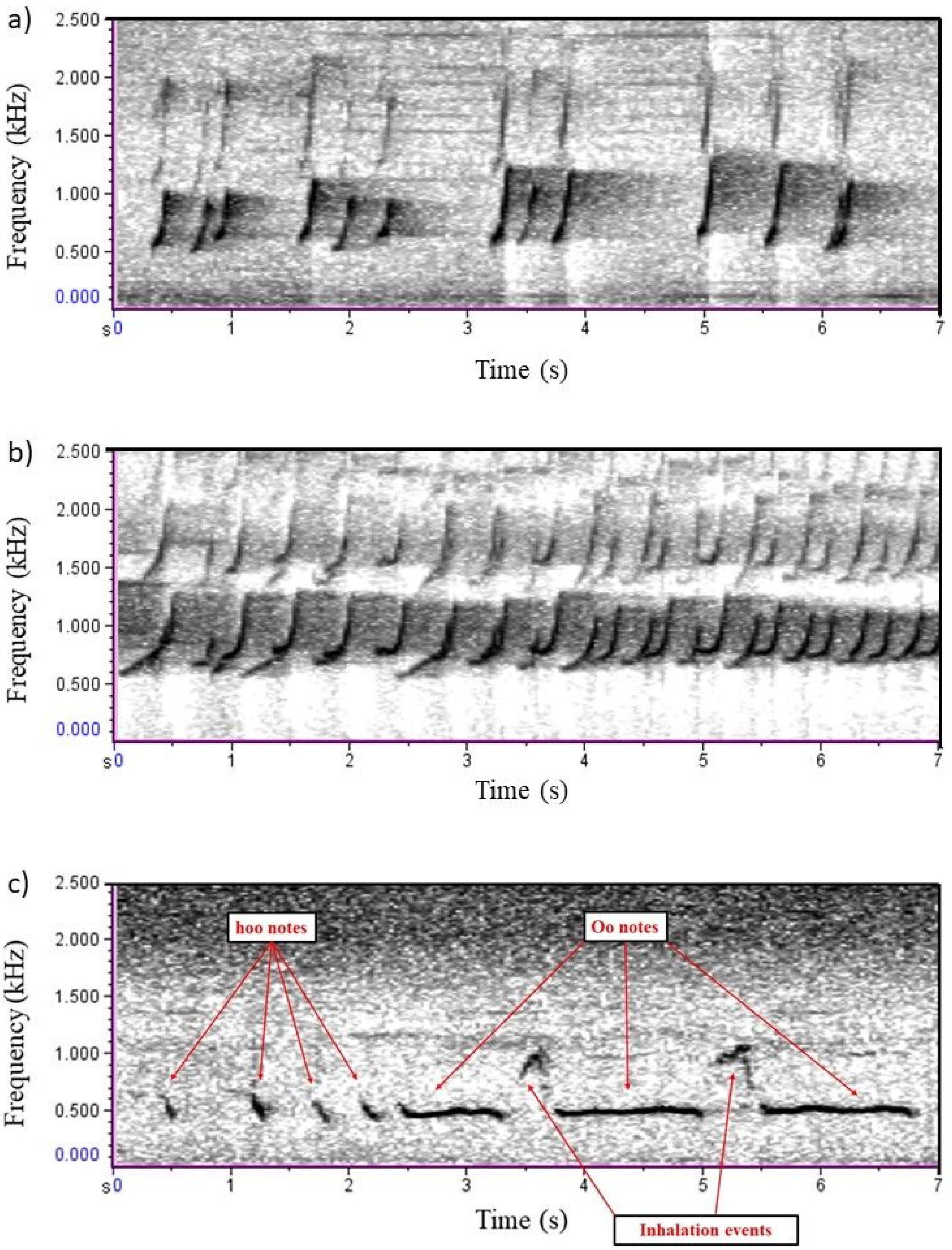

Figure 1.2. White-handed gibbon figures. a) Example of four clusters of male trills recorded during a duet. b) Example of an 'Ooaa burst' figure recorded during an 'Ooaa' duet produced by the mated pair of a gibbon group. The male starts the first "Ooaa" note, then the female joins with a quiet "Ooaa" until they synchronise their contribution almost perfectly. c) Example of 'Hoo-oo' figure recorded during a predator song. In the spectrogram, the exhaling notes are 'oo' notes followed inhaling events.

b. Figures

From their note repertoire, gibbons can assemble notes into specific segments to form acoustically and visually highly recognisable *figures*. To date, three classes of figures have been identified: the '*male trill'*, the *'Ooaa burst'* and the '*Hoo-oo*' figures (Raemaekers et al., 1984a).

Male trill (Figure 1.2.a): This figure is composed of a rapid repetition of 'wa' notes, each cluster being produced in one exhalation. This figure is male specific and can be uttered in the introductory sequence of duet, but mostly appears later in the song bout, in interlude sequences, to introduce a male structured phrase.

Ooaa bursts (Figure 1.2.b): This figure is mainly performed as part of 'Ooaa' duets, and only contains distinctive J-shaped 'ooaa' notes. 'Ooaa bursts' can also be produced sporadically in regular duets. It is structured as a rapid series of 'Ooaa' notes and can be uttered by the male only, in synchrony with the female and/or with offspring. This figure is acoustically very distinctive.

Hoo-oo figure (Figure 1.2.c): The Hoo-oo figure will be further discussed in the third chapter (Gibbon songs refer to external events). This quiet low-pitched figure is composed of both inhalation and exhalation events. The exhaling note can be either a short 'hoo' or a longer 'oo' note. The inhaling event is higher pitched.

c. Phrases

Gibbons have sex-specific structured sequences that are more complex than the figures described above, termed '*structured phrases'*. A phrase is defined as a single vocal event, produced by one individual, containing a series of notes, which are assembled together in a very specific sequence (Geissmann, 2002a). In white-handed gibbons, structured phrases are divided into three classes: the stereotyped female ***great call***, and two male specific structured phrases, the ***simple** and **quaver phrases*** (Raemaekers et al., 1984a), typically produced during male solo bouts but also as a reply to their mate great call.

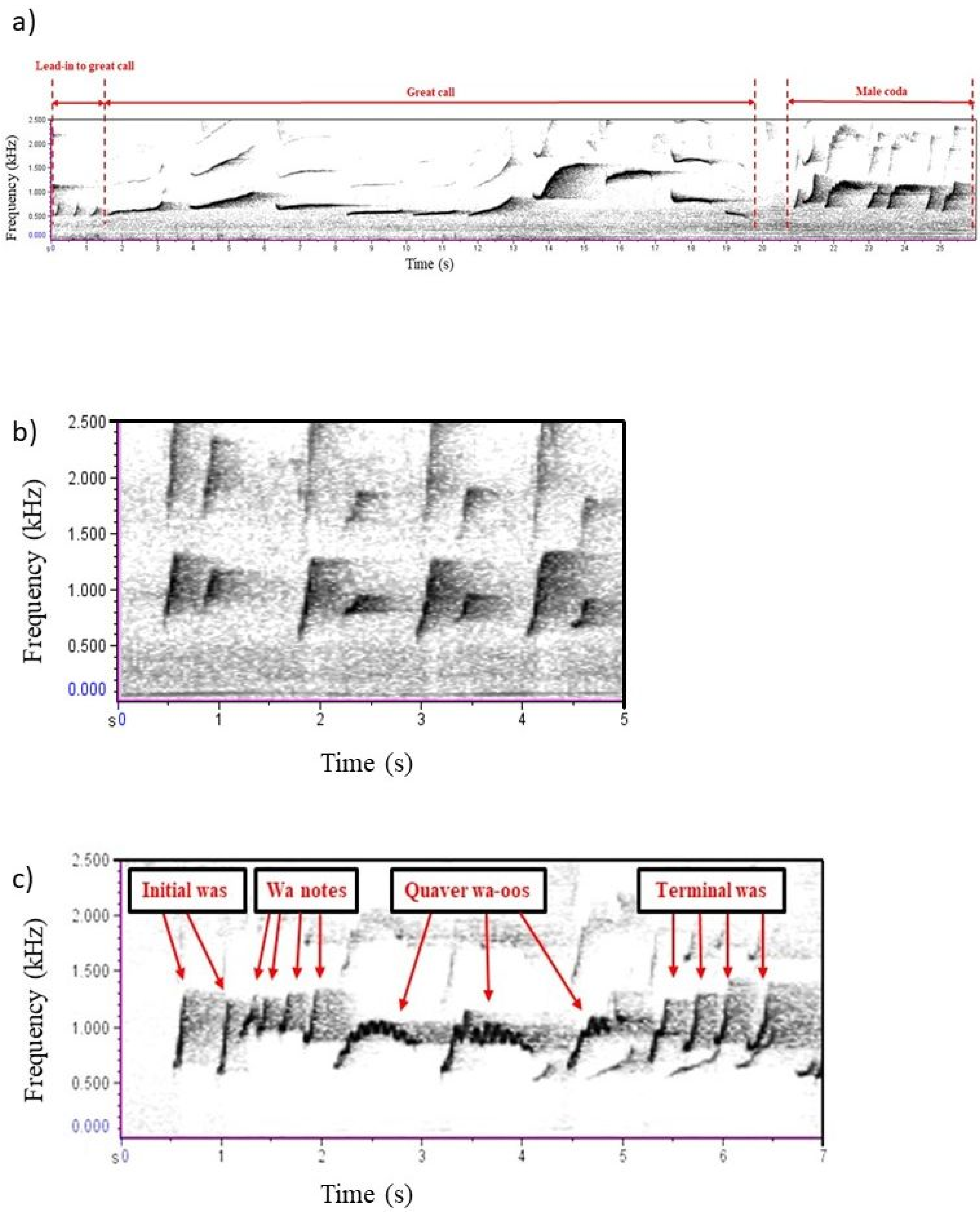

Figure 1.3. White-handed gibbon structured phrases. a) The female great call followed by her mate answering coda, which forms together the 'great call sequence'. This regular great call sequence was extracted from a duet produced by group M. A great call usually starts with the female 'lead-in to great-call' phase, followed by the so-called female great call. b) Male simple phrase extracted from the male coda of group E during a duet. c) Typical male quaver phrase, extracted from a duet performed by group W, as part of the interlude sequence.

The female great call (Figure 1.3.a) is the most stereotyped, recognizable and female specific phrase of all gibbon species; and in the white-handed gibbon's repertoire, the great call is the longest acoustic structure. The female produces great calls during morning duets, Ooaa duets, disturbed song bouts and predator songs. They can also be produced independently as solo sequences, during intergroup encounters or when the female and its mate are kept apart (Raemaekers et al., 1984a). A great call begins with a 'lead-in to great call' phase, composed of "hoo" notes produced by the female. During this sequence the male usually stops singing, letting the female uttering her great call (Raemaekers et al., 1984a). A great call lasts in average from 12 to 22 seconds and contains from 6 to 13 notes that increase in frequency until an initial first peak (from 400 to 1200 Hz), then decrease to form a "plateau" of several notes (from 400 to 600 Hz) to increase in frequency again until the climax note (from 1300 to 1600 Hz) and then fall again with the post climax notes (until 400 Hz). The male often adds his contribution at the end of his mate great call, under the form of a male simple or quaver phrase, termed as 'male coda' when produced as a reply to its mate great call. The combination of the female great call and its male answering coda is termed the 'great call sequence' (Raemaekers et al., 1984a).

The conspicuous female great call was widely thought to function primarily as a territorial component used for same-sex exclusion, warning about the ownership of one's territory in addition to its pair bond function, as the male coda often accompanied each great call (Mitani, 1985a; Raemaekers and Raemaekers, 1985a; Cowlishaw, 1992; Geissmann and Orgeldinger, 2000; Terleph et al., 2015, 2016). Great calls have a low intra-individual variation, meaning that great calls produced by one female throughout a song bout are highly homogeneous; but they differ considerably between individuals, this characteristic could potentially allow receivers to discriminate between different callers (Terleph et al., 2015).

The male simple phrase (Figure 1.3.b) is the basic form of the male quaver phrase without the quaver pitch modulations and is classically produced in male solos or given as male coda after female great calls. The male simple phrase is composed of "wa" notes from narrow to broad pitch range. Nonetheless some notes of the male simple phrase can also be subject to quaver pitch modulations as the male vocal contribution evolves throughout the song bout (Raemaekers et al., 1984a).

The male quaver phrases (Figure 1.3.c) are more variable than the highly stereotyped great call, however, this distinctive structured phrase can be divided into four distinct note types, produced in a specific temporal arrangement (Raemaekers et al., 1984a):

Type I: Initial ‘wa’ notes, with a large frequency range.

Type II: ‘Wa’ notes with a thinner frequency range, averaging 1000 Hz, given as in male trill as a quick ‘wa’ notes series.

Type III: ‘Wa-oo’ notes subjected to quaver modulations.

Type IV: Terminal ‘wa’ notes, uttered less speedily than the type II male trill, with a wider frequency range.

1.2 From basic notes to elaborate songs

All gibbons species are recognised to produce stereotyped long and loud complex vocalisations known as “songs” (Haimoff, 1984a; Geissmann, 1993, 1995, 2002a). The definition of a song was first given by Thorpe, in 1961, while describing the behaviour of singing birds, and corresponds to a series of notes pure in tone, usually of more than one type, that are produced in succession and associated to form a noticeable sequence or pattern in time. Then, with the first studies on gibbons, the definition was adapted to the Hylobatid family according to their singing behaviour, standardizing acoustical and organisational terms for all gibbon species, and a song bout was defined as the totality of notes, from the first to the last note, uttered by one gibbon group with, defined arbitrary, no period of silence of more than ten minutes between notes (Haimoff, 1984a).

a. The duetting mated pair: duet structure and functions

Duet song structure

White-handed gibbons are territorial, arboreal and mostly monogamous primates, living in the forest of South-East Asia. Gibbon groups comprise between two and six individuals, composed of the mated pair, and sometimes a secondary male and offspring. The adult pair routinely produces loud duet songs that can usually be heard by individuals occupying a neighbouring territories over distances of more than one kilometre, despite dense forest vegetation (Marshall and Marshall, 1976). Duet songs are given mostly mid-morning, peaking around 08:30 local time and are given anywhere in the territory (Raemaekers and Raemaekers, 1985a). Groups can be either single- or multi-males, but only the primary male and the mated female duet together. During duet songs, the mated pairs fuse their vocal contributions in an organised way, to produce complex, but relatively stereotyped alternated vocal interactions (Marshall and Marshall, 1976; Marler et al., 1977; Haimoff, 1983, 1984a; Marshall and Sugardjito, 1986). The combination of their vocal utterances during duets is thought to be subjected to social influences, with a higher degree synchrony between well- established pairs compared to newly formed pairs, such as a better anticipation of males to fall silent at their mate great call onset, facilitating great call transmission (Geissmann, 1999a; Terleph et al., 2017).

In nearly all duetting gibbon species, the duet song bout contains distinct sequences that are successively produced. The duet song is divided into three types of sequence: the ***introductory sequence***, the ***great call sequence*** and the ***interlude sequence***.

The introductory sequence (Figure 1.4.b) is only produced once, opens the duet and comprises all the notes uttered before the production of the first great call. The sequence lasts on average two minutes and is usually initiated by the male (Raemaekers et al., 1984a; Clarke et al., 2006; Andrieu et al. submitted a). This sequence is introduced by the production of soft “hoo” notes series (Figure 1.4.a) followed by to the production of “oo”, “wa”, “leaning wa” and “wa-oo” notes by the pair until the first great calls sequence (Figure 1.4.a, Raemaekers et al., 1984a).

The great call sequence (Figure 1.3.a) comprises the female’s great call usually followed by her mate’s answering coda. This sequence is produced repetitively, in a more or less rhythmical manner, throughout the duet song bout (Raemaekers et al., 1984a).

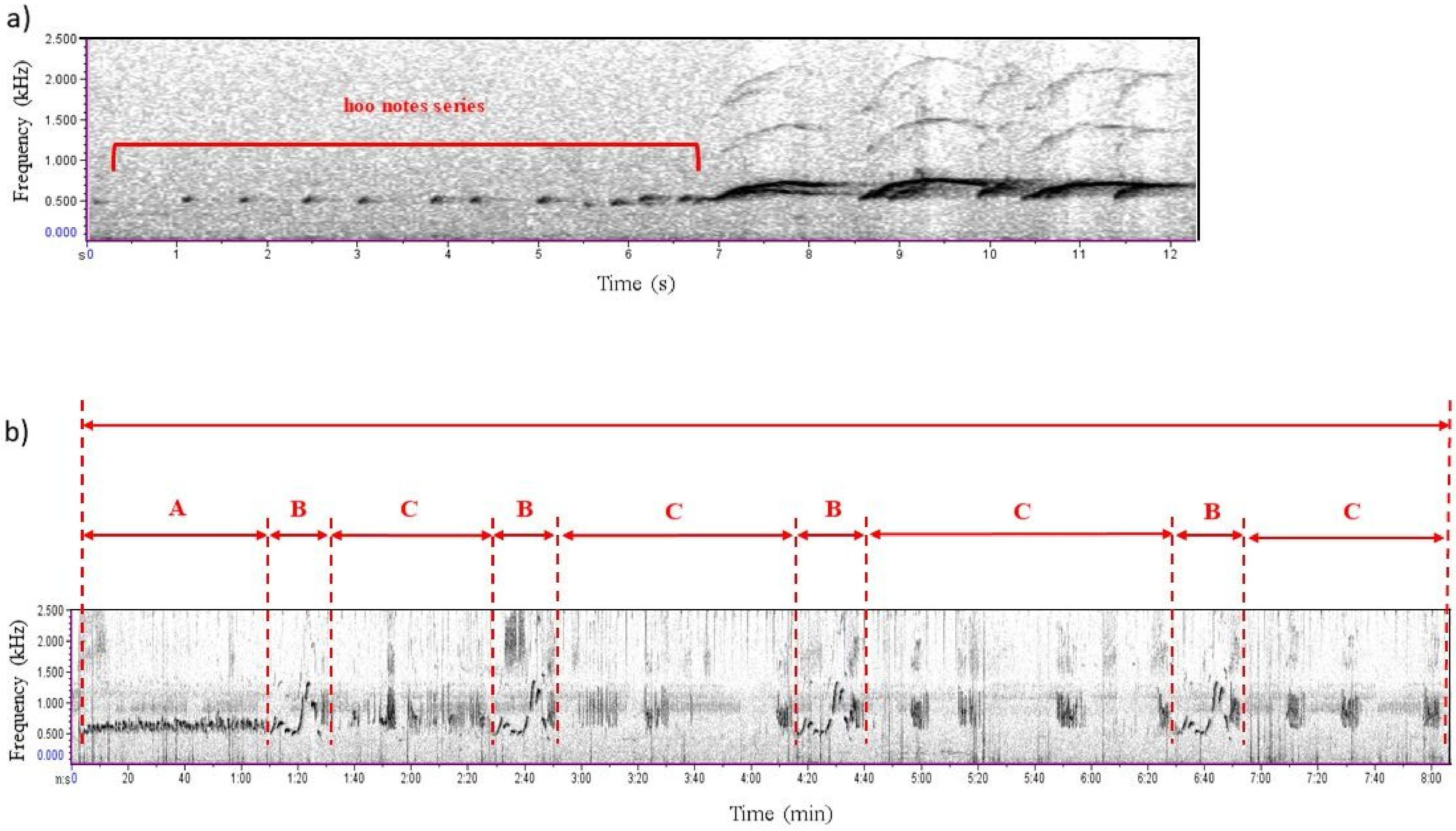

Figure 1.4. White-handed gibbon's duet song structure. a) Soft 'hoo' notes series opening duets. b) Duet song sequential structure (A: Introductory sequence, B: Great-call sequence, C: Interlude sequence).

The interlude sequence (Raemaekers et al., 1984a, Figure 1.4.b) includes all notes produced in between great call sequences and after the last great call sequence to the end of the song, and is the most variable of the three duet sequences. The male mostly contributes to this sequence compared to females, and sometimes the latter do not contribute at all. This sequence is composed of phrases separated by period of silence.

Duet songs generally open with the introductory sequence (A) that stops with the production of the first great call sequence (B) which is followed by the interlude sequence (C). The structure of a duet can thus be depicted as: ABCBCBCBC… The number of "BC" cycle is variable and the duet song bout can end up either on a great call sequence (B) or on an interlude sequence (C) (Raemaekers et al., 1984a).

Duet song function

Wickler (1980) considered that duet songs originate from joint territorial displays; the main function in most duetting species. Gibbons duet songs are audible up to 1 km, loud enough to be heard not only by their close neighbours but also by more distant groups (Marshall and Marshall, 1976; Raemaekers and Raemaekers, 1985b) are the primary method used by researchers to locate wild populations.

Duetting behaviour is known to function in territory defence and intergroup spacing, advertising about the ownership of one's territory to individuals that are part of neighbouring groups (Brockelman, 1975; Marshall and Marshall, 1976; Raemaekers and Raemaekers, 1985a). Territory borders comprise overlapping zone, which are the site of intergroup disputes where males are usually involved and actively defend borders by displaying, chasing and soft calling along their territorial borders, with females generally staying in the background (Carpenter, 1940; Ellefson, 1968; Gittins and Raemaekers, 1980; Raemaekers and Raemaekers, 1985a). Playback experiments revealed that vocal response rate to neighbours' duet songs and the intensity of the behavioural response associated (e.g. approaching the source of "conflict") were found to vary according to the location of the duet emitter (Chivers, 1974; Brockelman, 1975; Tenaza, 1976; Raemaekers and Raemaekers, 1985a, 1985b), suggesting that receivers use songs as a means to identify and locate out-group individuals, which can potentially mediate intergroup spacing (Mitani, 1985b).

Duet songs are also thought to function in pair-bonding, mate attraction and as a mean of same- or opposite-sex exclusion (Raemaekers and Raemaekers, 1985a). During playback experiments of 'stranger males' and 'strangers females' solos bout from the centre of territory of the gibbons groups, the individuals of the tested mated pairs approached systematically the speaker location but in different ways, females behaved more strongly to 'intruding females' and vice versa for the male (Raemaekers and Raemaekers, 1985a). Duet songs can also serve as a mean to identify individuals, social status and also a mean to assess physiological conditions. White-handed gibbons were observed to sometimes pause during their own duet to listen the great-call sequences of neighbours duet songs (Wickler, 1980; Raemaekers and Raemaekers, 1985b), this specific sequence was found to convey information about external events, physiological state and identity of callers (Terleph et al., 2015, 2016). Males were also found to have acoustic inter-individual differences (Barelli et al., 2013a). Besides, males androgen levels were found to be correlated with higher pitched call sequence production and social organisation in white-handed gibbons; with male living in pairs having higher androgen level than primary males living in multimale groups (Barelli and Heistermann, 2012; Barelli et al., 2013a).

b. Disturbed song bouts and predator songs

Apart from the remarkably loud and conspicuous duet songs performed by the mated pair during the morning hours, white-handed gibbons also produce songs to external disturbances and predators (Raemaekers et al., 1984a; Clarke et al., 2006).

Disturbed song bouts

This class of song bout was first labelled 'disturbed song bouts' instead of 'alarm' by Raemaekers and collaborators, (1984a), because they were induced not only by external stimuli but also intra and inter-group causes. This category contained disturbed song bouts in response to alarm calls by sympatric species, conflicts with groupmates or neighbouring groups during territorial disputes, other groups' disturbed song bouts, human presence and potential predators ("*…probably a snake*" (Raemaekers et al., 1984a).

The introductory sequence of disturbed song bouts was distinguished from those of duet song bouts by their lack of rhythm and the participation of other individuals besides the mating pair, going along with lower delivery rate of great calls and lasting longer than normal duets (Raemaekers et al., 1984a). The song bout begins with low frequency 'wa' notes, then gradually increasing in frequency, with the production of ‘oo’, ‘waoo’ and ‘sharp wow’ notes, and was found to be longer in duration than duet songs (N.B. plosive low frequency ‘wa’ notes were found to be part of another class of notes, termed ‘hoo’ notes ;Clarke et al., 2006).

Predator songs

In some gibbon species, the abilities of uttering alarm calls are just based on rare observations or description (Silvery gibbons (*Hylobates moloch*) (Geissmann and Nijman, 2006), Müller gibbons (*Hylobates muelleri*) (Mitani, 1985c), Crested gibbons (*Nomascus concolor*) (Fan et al., 2006)) and in others this vocal behaviour is still unknown (Pileated gibbons (*Hylobates pileatus*), Agile gibbons (*Hylobates agilis*), Siamangs (*Hylobates syndactylus*), or Yellow-cheecked gibbon (*Nomascus gabriellae*)).

So far, white-handed gibbons are part of the only species within the Hylobatid family that have been observed, described and experimentally tested to produce long and loud songs in response to potential predators, such as large cats and snakes. Predator songs are composed of the same song units as regular duet songs, although arranged in different ways (Clarke et al., 2006). Clarke and collaborators, (2006, 2012) have presented different life-size predator models to free-ranging individuals in Khao Yai National Park, Thailand. Crested serpent eagle (*Spilornis cheela)*, reticulated python (*Python reticulatus*), clouded leopard (*Neofelis nebulosa*) and tiger (*Panthera tigris*) models induced mixed vocal and behavioural responses in gibbons groups.

Large cats models elicited strong anti-predation behaviours and subjects reliably sang in response to them, whereas python and raptor models elicited respectively few singing (production of soft songs composed of ‘hoo’ and ‘oo’ notes) or no songs in response (but only few ‘hoo’ notes produced), associated with weak anti-predatory behaviours (Clarke et al., 2012). Even though predator models unequally affected gibbons, the songs produced in response to large cat models were found to be structurally different compared to regular duet songs. Predator songs are introduced by a longer series of soft ‘hoo’ notes, with a highly delayed great call production accompanied by early male responses, meaning that introductory sequences in predator songs last longer compared to duets. (Clarke et al., 2006). The total song bout duration is also longer than regular duets, and contain more ‘sharp wow’ notes, which production is also delayed compared to regular duets. (Clarke et al., 2006).

Moreover, gibbons also produce another different structural predator song type in response to the reticulated python model presentations (Clarke et al., 2006), only composed of soft ‘hoos’ and ‘oo’ notes. In another study, Clarke and collaborators, (2015) found additional acoustic features concerning the opening soft ‘hoo’ notes that differed depending on the external context, with frequency related ‘hoo’ notes measures discriminating ‘hoos’ emitted for big cats and those produced in duets, as well as longer inter-call interval between ‘hoo’ notes that are different in raptor ‘hoos’ than for duet and big cat ‘hoos’ (Clarke et al., 2015).

c. Other songs

White-handed gibbons also produce other type of vocalisations, such as male and female solos, and another duet type, the ‘Ooaa’ duet.

Ooaa duets are the most enigmatic duet form in white-handed gibbons. The structure of this duet type can be similar to regular duets (i.e. introductory, great call and interlude sequence), but differs in containing more or exclusively ‘Ooaa burst’ figures in introductory and interlude sequences (Figure 1.2.b) or can even only contain ‘Oaaa bursts’ without any duet-like sequential arrangement (Raemaekers et al., 1984a).

These duets are rarer, longer than regular duets and highly infectious, so that they ‘contaminate’ rapidly neighbouring groups in wave of ‘Oaaa’ duets singing (Raemaekers and Raemaekers, 1984b). So far, despite the infectious spreading to neighbouring groups, naturally and experimentally observed, we still do not know how this singing form is elicited, its function and what meaning it can potentially convey (Raemaekers and Raemaekers, 1984b).

Male solo songs are produced by primary, secondary and subadult males and are typically the longest song bouts among all classes produced by gibbons, and composed low ‘oo’ notes that develop into male simple and quaver phrases (Figure 1.3.b and Figure 1.3.c), with additional ‘waoo’ notes (Raemaekers et al., 1984a). They are produced during the morning hours, from 2 hours before dawn from their night tree to dawn, but also later in the morning hours and are hypothesized to function in territory advertisement (Raemaekers et al., 1984a).

Female ‘solo songs’ were described as simply composed of an isolated great call, sometimes repeated and occasionally produced with some ‘wa’ notes; they are uttered as contact call bouts, when females are separated from their mate, or when the latter is involved further away in territorial disputes (Raemaekers et al., 1984a). In the current literature, the status of proper solo song bouts was abandoned, due to their isolated nature of production, and surely their lack of complexity compared to other hylobatids species (Geissmann, 2002a; Reichard et al., 2016). For instance, female Silvery gibbons (*Hylobates moloch*) produce solo song bouts that are sequentially similar to duet songs produced by other duetting species mated pairs, composed by an introductory, great call and interlude sequences (Geissmann and Nijman, 2006).

1.3 Thesis research

a. Study site and subjects

Study site

My observations and experiments on white-handed gibbons were conducted in Khao Yai National Park, Thailand (101°22'E, 14°26'N), 130 km North-East of Bangkok. Khao Yai National Park was established in 1962 and covers a total area of 2,165 km^2. The Mo-Singto – Klong E-Tau research site is located in the central portion of the park at 730-870m above sea level, receiving over 3,000 mm of rain each year, with a marked wet season from May through October and relatively dry conditions from November until April (Whitington and Treesucon, 1991).

<u>Subjects</u>

The study population of Mo Singto - Klong E-Tau area consisted in 13 territorial groups gibbons (Figure 1.5), eight monogamous groups and five polyandric groups, totalling 53 individuals (Table 1.2). In polyandric groups, the primary male is defined as the one being engaged in duetting and copulation with the adult female, whereas the secondary male is not supposed to engage copulating nor duetting with the adult female. Some observations showed that these secondary (unrelated) males sometimes copulate with the female, but this happens significantly less often than copulations by the primary male (Reichard, 1995; Barelli et al., 2013b). White-handed gibbons live in nuclear family units of two to six individuals (Table 1.2), in home ranges of between 400 and 500 metres in diameter, ranging from 2.7 to 27.5 ha (Reichard and Sommer, 1997; Andrieu et al. submitted a, Figure 1.5).

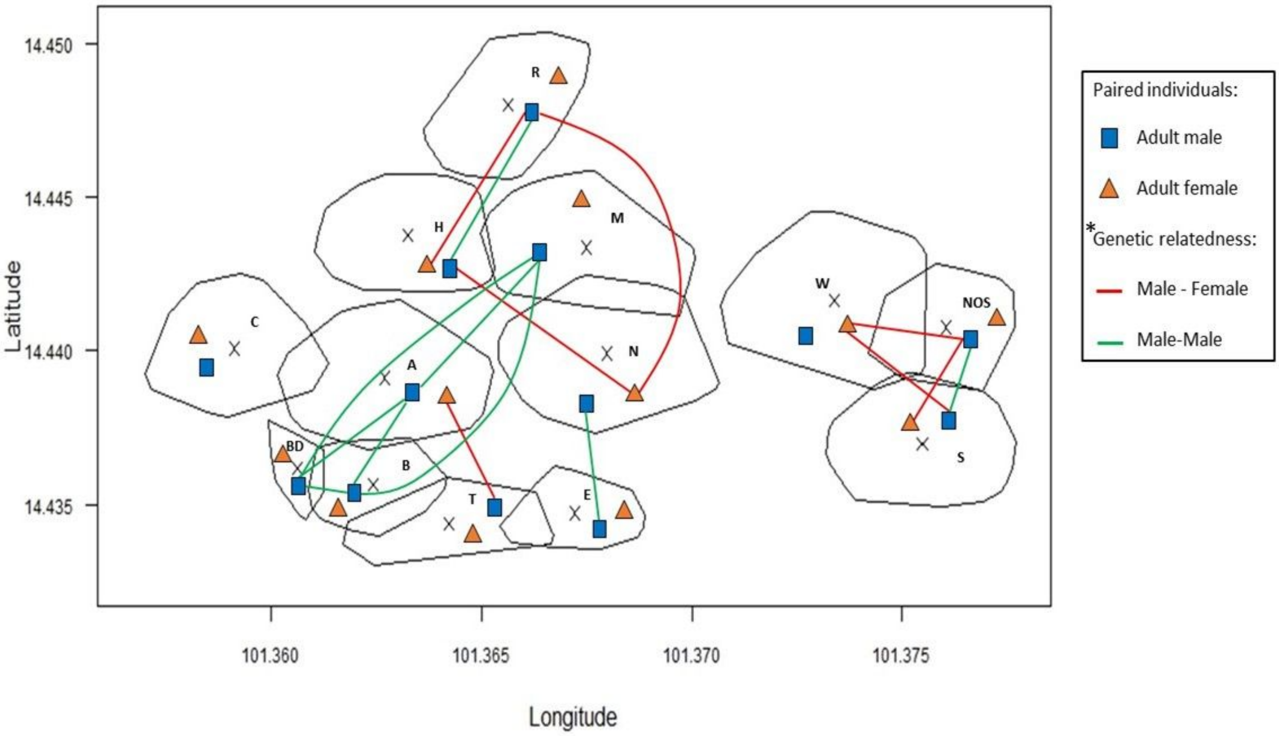

Figure 1.5. Home ranges of 13 habituated white-handed gibbon groups of Mo Singto - Klong E-Tau area, Khao Yai National Park, Thailand (Groups: A, B, BD, C, E, H, M, N, NOS, R, S, T, W) and genetic relationships amongst paired adult males and females. The map was drawn using 95% kernel density estimation (KDE). X=home range centroid. The genetic relationships between individuals were updated and redrawn from (Matsudaira et al., 2018). Red lines represent the presence of genetic relationships between paired females and paired males and green lines represent the presence of genetic relationships between paired males paired amongst the different study groups.
* Kinship: parent-offspring / full-siblings / half siblings.

The habituation of gibbons in the Mo Singto - Klong E-Tau area started in the 1980s, their population density is estimated to 15.9 individuals per km^2 (Barelli et al., 2007). Long-term demographic data were collected through *ad libitum* observations while passively following gibbons during their daily foraging course (Reichard and Barelli, 2008; Reichard, 2009; Reichard et al., 2012). Genetic relationships were determined in a previous study by Matsudaira et al., 2018 (see Chapter 2. Nepotism enhances intergroup communication in gibbons).

Table 1.2 Group compositions of the white-handed gibbon population at the Mo Singto-Klong E-Tau study site (August 2014)

Group	N individuals	Identities and age/sex class
A*	3	2AM: **Chuu** and Chokdee, 1AF: Andromeda
B*	5	2AM: **Chet** and Cassius II, 1AF: Baak, 1JF: Baan, 1I?
BD	3	1AM: **Christophe**r, 1AF: Bodana, 1I?
C	3	1AM: **Chao**, 1 AF: Chinda, 1I?
E	3	1AM: **Noi**, 1AF: Elane, 1JM?
H	4	1AM: **Felix**, 1AF: Hannah, 1JF: Hera, 1I?
M	5	1AM: **Chikyu**, 1AF: Rung, 1SAF: Malai, 1JM: Muang, 1I?
N*	6	2AM: **Nithat** and Claude, 1AF: Hima, 1SAM: Naam, 1JF: Nan, 1I?
NOS*	5	2AM: **San** and Ness, 1AF: Nasima, 1J?: Nishan, 1I?
R	4	1AM: **Henry**, 1AF: Brit, 1AF: Rarin, 1I?
S	3	1AM: **Samran**, 1AF: Sofi, 1JM: Samri
T	5	1AM: **Amadeus**, 1AF: Brenda, 1SAM: Teera, 1JM: Toona, 1I?
W*	4	2AM: **William** and Wotan, 1AF: Sari, 1I?

*Multi-male group, primary male in bold; M=male, F=female, ?=sex unknown, A=Adult (age >8years), SA= Sub-adult (5-8 years), J=juvenile (2-5 years), I=infant (<2years).

b. White handed gibbon: taxonomy and song evolution

Hylobatids or small Asian apes, comprising gibbons and siamangs, occupy the most distant branch of the Hominoids family, from which they diverged seemingly in Africa into the early (26 mya) to early middle Miocene (15 mya) according to molecular clock estimates, and then spread to Eurasian forests (Matsui et al., 2009; Chan et al., 2010; Matsudaira and Ishida, 2010; Carbone et al., 2014; Reichard et al., 2016). But it is started to be commonly accepted

Table 1.3 Song types produced by the different Hylobatidae species

Genus	Species	Song types				
		Female solo	**Male solo**	**Duet song**	**Alarm calls**	**Source**
Hylobates	*H. moloch*	yes	yes (rare)	No	yes	a, b, c
	H. klossi	yes	yes	No	yes	d, e
	H. agilis	no	yes	yes	ND	f, g
	H. muelleri	yes	yes	yes	yes	h, i, j
	H. pileatus	no	yes	yes	ND	k, l
	H. albibarbis	no	yes	yes	ND	m, n
	H. lar	**no**	**yes**	**yes**	**yes**	**o, p**
	H. funereus#	no	yes	yes	ND	q
	H. abbotti#	ND	ND	ND	ND	
Hoolock	*H. hoolock*	no	no	yes	yes	r
	H. leuconedys	ND	ND	yes	ND	s, t
Nomascus	*N. concolor*	no	yes	yes	yes	k, u
	N. hainanus	no	yes	yes	yes	v, w
	N. nasutus	no	yes	yes	ND	k, x
	N. gabriellae	no	yes	yes	ND	k, y
	N. siki	ND	yes	yes	ND	k, y
	N. leucogenys	yes*	yes*	yes	ND	k, y
	N. annamensis	no	yes	yes	ND	z, α
Symphalangus	*S. syndactylus*	no	no	yes	contradictory	e, k, β, γ

Yes, present; no, absent; ND, no data available; *when adult individuals are unpaired, **#sub-species.** [a]Geissmann and Nijman (2006), [b]Geissmann et al. (2005), [c]Kappeler (1981), [d]Dooley et al. (2013), [e] Tenaza and Tilson (1977), [f] Haimoff (1984), [g]Haimoff and Gittins (1985), [h]Inoue et al. (2013), [i]Mitani (1985a), [j]Mitani (1985b), [k]Geissmann (2002), [l]Traeholt et al. (2006), [m]Wanelik et al. (2013), [n]Cheyne et al. (2007), [o]Raemaekers et al. (1984a), [p]Clarke et al. (2006), [q]Inoue et al. (2016), [r]Kumar et al. (2013), [s]Fan et al. (2013), [t]Yin et al. (2016), [u]Fan et al. (2006), [v]Deng et al. (2014, 2016), [w]Haimoff (1984b), [x]Feng et al. (2014), [y]Konrad and Geissmann (2006), [z]Williams (2017), [α]Van Thien et al. (2017), [β]Orgeldinger (1997), [γ]Morino (2010).

that they split from hominoids around 16.26 mya (Thinh et al., 2010a). As the oldest known hylobatid fossile, the *Yuanmoupithecus* found in China, was dated back to approximately 8 mya, there is a huge gap in their geographic spreading and distribution history, as well as in the evolutionary adaptation of small Asian apes from their splitting time from Hominoids to their first know presence in Asia (Harrison, 2016). Hylobatids, then, radiate rapidly into four different genera between 8.34 to 6.69 mya with one of the latest to have emerged being the *Hylobates genera* (Thinh et al., 2010a) comprising white-handed gibbons. Interestingly, the emergence of four hominins genera, including the *Homo* genus, concurrently emerge at a similar period and with a comparable time span radiation, making hylobatids a species of interest to compare the dynamic evolutionary process of hominin evolution (Zichello, 2018).

The current mostly accepted classification divides Hylobatids into four monophyletic groups, considering the number of their diploid chromosomes (*Hylobates*: 2n=44*, Hoolock*: 2n=38*, Nomascus*: 2n=52 and *Symphalangus*: 2n=50), as well as their acoustic singularities, and their morphological and molecular differences (Schultz, 1933; Boer, 1972; Chiarelli, 1975; Prouty et al., 1983; Mootnick and Groves, 2005; Roos, 2016). To date, 17 confirmed species of Hylobatids are recognised with the recent discovery of a new (sub-)specie, the Skywalker hoolock (*Hoolock tianxing*) in 2017 (Geissmann, 1995, 2002a, 2002b; Mootnick and Groves, 2005; Walker et al., 2007; Van Ngoc Thinh et al., 2010b; Mittermeier et al., 2013; Roos, 2016; Chan et al., 2017; Fan et al., 2017;) (Table 1.3). Though, there is still debate on the classification of *Hoolock tianxing*, *Hylobates abotti* and *Hylobates funereus* as sub-species or species (Geissmann, 1993, 1995, 2007; Groves, 2001; Thinh et al., 2010a, 2010c, 2010; Mootnick and Fan, 2011; Anandam et al., 2013), and so I will simply consider them here as *Hoolock* and *Hylobates* sub-species.

Gibbon songs are species-specific and, to a certain extent, sex-specific (Haimoff, 1984a; Marshall and Marshall, 1976; Geissmann, 1993, 1995). Most Hylobatid species are recognised to produce long, loud and conspicuous songs including duets. Gibbon duetting behaviour is by far the most studied component of their vocal behaviour, with many open questions on how this behaviour has emerged evolutionarily. All but two gibbon species of the four different genera produce duets. One unresolved issue is whether the absence or the presence of duetting is the primitive state. So far, duetting was commonly accepted to be the Hylobatids' ancestral state with the two non-duetting species the Silvery gibbons and Kloss's gibbons (*Hylobates moloch* and *Hylobates klossi*), having lost their duetting faculties secondarily as a synapomorphic trait (Geissmann, 2002a, 2002b).

Though, the first resulting attempts to reconstruct hylobatids phylogeny were contradictory and mainly based on vocal, anatomical and morphological traits, in accordance with the results of the earliest molecular procedures, and none showed identical results, with the two non-duetting species not having each time diverged in the same branch. It was thought to be explained by the short evolutionary period in which all extant gibbons lineage diverged (Thinh et al., 2010a). Many studies, including molecular ones, focussed on resolving the conflicting phylogenetic relationships between the four genera of the Hylobatidae family, and indeed, the different genera may not have diverged in a strict branching pattern (Mootnick and Groves, 2005). It is now commonly accepted that duetting was the ancestral form of hylobatids, and that the geographical isolation of the non-duetting species from other duetting species could well have led to the loss of their duetting abilities as a synapomorphic trait, more likely resulting from a convergent evolution of their duet loss than inherited from an ancestral form. Silvery gibbons are found in the western part of Java and Kloss's gibbons in the Mentawai islands and both share no borders with any of the duetting species, contrary to the rest of the duetting hylobatids that are at least in contact with another duetting species (Chivers, 1977; Chivers and Gittins, 1978; Marshall and Sugardjito, 1986). It might be indeed plausible that silvery and Kloss's gibbons, might have derived later from a duetting ancestor, leaving behind their duetting behaviour, with the male and female splitting off their vocal contributions to specialise only in solo song bouts. For example, in silvery gibbon, it was shown that males literally avoid singing with mates and that they even interrupt their ongoing vocal production if females start singing (Geissmann, 2000a), as male totally withdraw from duetting by avoiding singing, giving away its contribution to the female only. Additionally, the structure of silvery gibbon female solos, composed of the tree same type of sequence as duet songs (i.e. introductory, great call and interlude sequences), strengthens the hypothesis, that this species could have evolved from a duetting ancestor and derived later from the hylobatid tree, losing or temporally displacing the male part of the duet (Geissmann, 2002a).

As mentioned earlier, the conspicuous duet song of gibbons has been given attention for many decades, and white-handed gibbons songs (i.e. duet songs and male solo songs) have been extensively studied, regarding their respective function and acoustic structure (Raemaekers et al., 1984a; Raemaekers and Raemaekers, 1985b; Terleph et al., 2015, 2016, 2017). But not all hylobatids species have been studied equally regarding their song structure and production skills, and compared to other species, for which we have little or no knowledge about their faculty to produce alarm songs in response to predators, human presence or other specific disturbance. In these other species, the capacity to utter alarm songs was sometimes documented as direct observations but never systematically studied (Table 1.3). White-handed gibbons represent the only species that has been systematically tested for their capacity to produce predator songs. Their sequential structure was found to differ from regular songs (Clarke et al., 2006, 2015), making them a promising species to study coding abilities in non-human primates and to explore the evolutionary roots of human language within the primate phylogeny.

c. Research project outline

Background

Many animal species, possess long-range communication signals that function to attract mates, defend resources or dissuade rivals, but how such functions are mechanistically implemented has typically remained unclear. One common strategy is to produce a long-distance signal to provide out-group members with information about the signaller itself, such as its spatial location, reproductive quality, general competitive ability, current motivation to defend a resource, or to predator encounters. Nevertheless, determining the function and specific meaning of vocalizations has often remained a considerable challenge.

White-handed gibbons have a rich and structurally complex vocal repertoire that contain a diverse range of song notes, assembled into more complex vocal utterances, such as ‘figures’ and ‘phrases’, later combined into loud and complex songs. Due to the numerous combinatory possibilities, callers can produce a range of structurally different song types (i.e. duet and predator songs), that could potentially convey specific information that could be meaningful and benefit neighbouring groups. Previous studies and natural observations revealed that groups sometimes reply by the matching song type when hearing other groups singing duet or predator songs (Clarke et al., 2006; Raemaekers et al., 1984a). However, the hypothesis that individuals can discriminate these different vocalisation types has never been experimentally tested. The structural differences found between duet and predator songs were mainly based on comparative studies between spontaneous duet songs and songs elicited by predator model presentations of a clouded leopard (Clarke et al., 2006). While groups were also observed to sing to reticulated pythons (see Chapter 4. Gibbon songs refer to external events), their capacity to encode different predator types has never been addressed.

Their conspicuous duet songs are also of particular interest for several other reasons. First, duetting evolved independently in a large number of taxa, including insects, anurans, birds, mammals and humans (Thorpe et al., 1972; Farabaugh, 1982; Haimoff, 1986; Nietsch, 1999; Tobias et al., 1998; Bailey and Hammond, 2003), suggesting a strong evolutive selection for this behaviour (Von Helversen, 1980). In Insects, duetting behaviour was found to have evolved several times independently. The same process was shown in birds (Thorpe et al., 1972; Farabaugh, 1982; Bailey, 2003; Hall, 2004, 2009) and non-human primates in which duetting is thought to have evolved independently at least four times (Haimoff, 1986; Geissmann, 2000a, 2000b). Duet singing mainly develops in monogamous and territorial species that live in dense tropical forests (Thorpe, 1961; Thorpe et al., 1972; Wilson, 1975) and among non-human primates, only species belonging to four different families are recognised to produce duet songs (Figure 1.6).

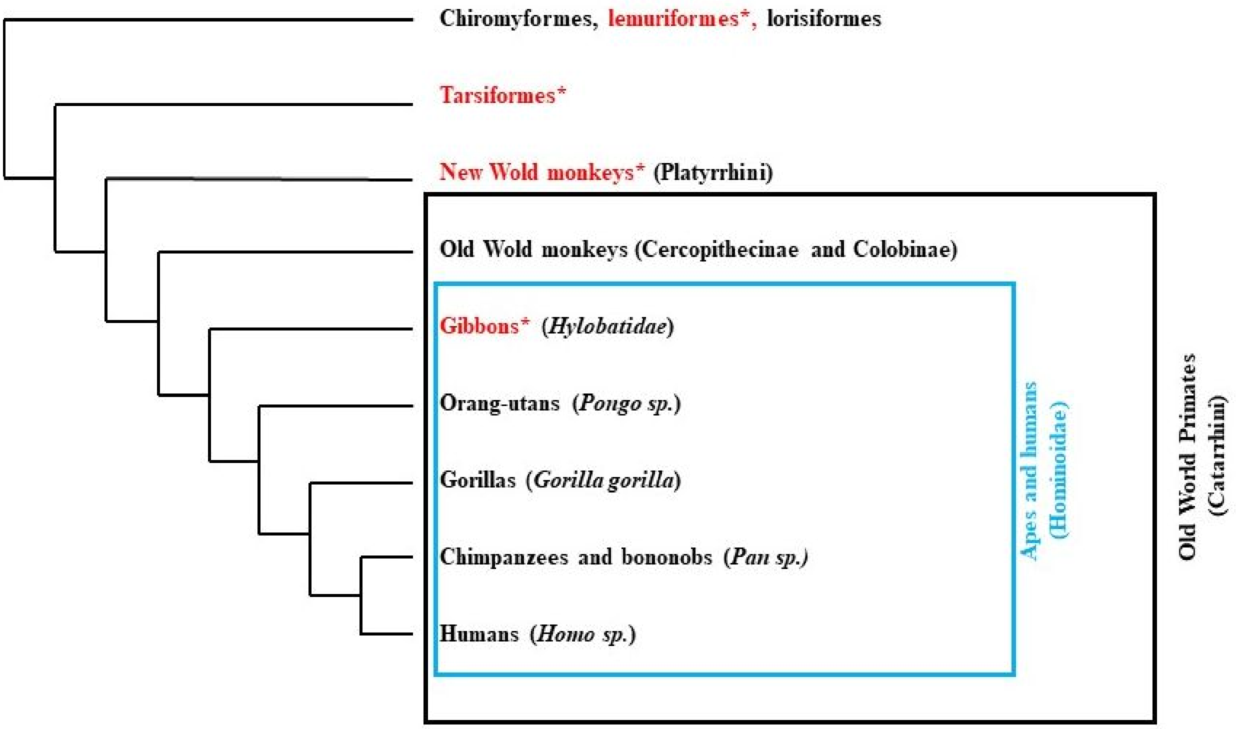

Figure 1.6. Phylogeny of the primate order. Figure was redrawn and adapted from (Pozzi et al., 2014). * represents non-human primates families where duetting species have been identified.

The Milne Edwards' sportive lemur regularly engages in duet calling behaviour with an equal participation in the vocal contributions of male and female (Méndez Cardenas and Zimmermann, 2009). Spectral tarsier (*Tarsius spectrum*) and Dian's tarsier (*Tarsius dianae*) give morning duets which function in territorial advertisement and as way to congregate at sleeping sites (MacKinnon and MacKinnon, 1980; Nietsch, 1999). Titi monkeys duet mostly during territorial encounters with neighbours which is thought to mediate intragroup and intergroup spacing (Robinson, 1981; Müller and Anzenberger, 2002). And among Hominoidae, the only non-human primate family adapted to duet performing are Hylobatids, and then Humans.

Secondly, such long-range communication signals typically encompass the average spacing of individuals, such as the modality of signal exchanges can be subject to interference, (McGregor, 2005), meaning that individuals within the audibility range can interfere with an ongoing signal, by overlapping it and totally masking it, or rendering its propagation difficult. However, individuals can also temporally alternate their signals by respecting each other's turn, referred as turn-taking, as in spoken language (Sacks et al., 1974; Yoshida and Okanoya, 2005), or by temporally overlapping it. In white-handed gibbons, both modality of response can be found during long-range duet morning exchanges with neighbours (Raemaekers and Raemaekers, 1985b). The investigation of duet interactive singing in gibbons (i.e. turn-taking vs counter-singing), especially the mechanisms underlying duet singing exchanges modalities is particularly relevant topic regarding several current research question or theories: the evolution of acoustic turn-taking abilities in compliance with the evolution of human communication system with the study of a phylogenetically closer species to Humans; and the value of signal or song overlap (i.e. counter-singing) which is controversial among Bird scientists.

Whereas signal overlap in birds is argued to be a mere by-product of chance (Searcy and Beecher, 2009), other researchers demonstrated that counter-singing could function in signal output amplification by conjointly uttering vocalisations (Narins, 1992; Greenfield, 1994b), as a male quality signal being an honest indicator of male quality and condition (Gil and Gahr, 2002), but mainly an agonistic or submissive signal to escalate or de-escalate aggressive interactions (Kunc et al., 2006; Naguib and Mennill, 2010). Most bird studies showed that individuals adapt their behaviour in response to overlap (Dabelsteen et al., 1997; Naguib, 1999; McGregor et al., 2000; Todt and Naguib, 2000). Given the territorial and mate defence and attraction functions allocated to gibbon duet songs, exploring whether groups can adapt their singing behaviour to counter-singing opponents, or whether other factors can have an effect of the probability of group to engage or not in counter-singing duet exchanges, could make them an alternative species to complement bird studies.

White-handed gibbons were also found to engage in duet turn-taking exchanges with neighbouring groups (Raemaekers and Raemaekers, 1985b). In New World monkeys, marmosets (*Callithrix jacchus*) exchange vocalisations in alternation without overlapping conspecifics (Takahashi et al., 2013) and squirrel monkeys respect an order in vocalising depending on the recognition of other individuals' chucks (Symmes and Biben, 1988). In Old World monkeys, turn-taking rules were found to be cognitively relevant for adult Campbell's monkeys (*Cercopithecus campbelli*) (Lemasson et al., 2011), and to date this ability was also shown in a closer phylogenetic group to Humans, Hylobatids such as Siamang (*Hylobates syndactylus*) (Geissmann and Orgeldinger, 1997), Mueller's gibbons (*Hylobates muelleri*) (Inoue et al., 2013) and White-handed gibbons (Raemaekers and Raemaekers, 1985b). In apes, the only species that were found to utter turn-taking vocalisations are gibbons and then Humans whereas Orang-utans (*Pongo sp.*), Bonobos (*Pan paniscus*) and Chimpanzees (*Pan troglodytes*) were found to prioritize gestural turn-taking at the expense vocal turn-taking (Call and Tomasello, 2007; Rossano, 2013; Rossano and Liebal, 2014). Taking turns to communicate is an essential prerequisite in human communication, and recently this ability has received considerable attention in animal interactive acoustic exchanges research. Hylobatids occupy the most basal branch of the hominoid family, and the investigation of their song production capacity could help us provide new insights about the evolutionary mechanisms of complex communication systems, such as speech and language.

Research projects

To explore the mechanisms of the flexibility of their vocal behaviour and assess their ability to code their songs depending on the context, my whole research was conducted in Khao Yai National Park in Thailand, where 13 groups on wild gibbons were available for the study.

In a first project, I investigated the factors underlying duet interactive singing between white-handed gibbons' groups. The mated pair produces highly synchronised duet songs, which travel over long distances up to one kilometre (Marshall and Marshall, 1976), suggesting that they are aimed at neighbouring individuals (Mitani, 1985b). Although some duet songs are produced spontaneously, duet singing is often a response to previous songs by another group, which can travel as waves of singing throughout a large part of the forest. In previous a previous study, playback experiments revealed that gibbon vocal and behavioural response rates tend to vary according to the location and distance of the interacting groups (Raemaekers and Raemaekers, 1985a), suggesting that receivers use songs as a means to locate neighbouring groups. Providing locational information is mutually beneficial as it allows groups to avoid conflict by adjusting their home range use (*Hylobates lar and Symphalangus syndactylus*: Carpenter, 1940; Chivers, 1974; Ellefson, 1968; Brockelman and Srikosamatara, 1984; Sommer and Reichard, 2000; *Hylobates muelleri:* Mitani, 1985c). This type of signalling is assumed to be honest (Kitchen et al., 2009), enabling receivers to take adaptive decisions and avoid unnecessary and costly interactions. Another strategy is to use long-distance signals as a means of engaging in intergroup conflict, such as to claim space or to access a resource (McGregor, 2005).

In white-handed gibbons, duet responses are more likely given to the songs of neighbouring than non-neighbouring groups and, generally, groups avoid overlapping with ongoing songs produced by other groups, which leads to turn-taking exchanges (Raemaekers and Raemaekers, 1985b). Sometimes, however, songs are produced such that they overlap with already ongoing singing by another group, which has been defined as counter-singing. Thus, I was particularly interested in the mechanisms that could drive gibbon groups to whether engage in competitive counter-singing duet exchanges (i.e. defined as one group starting their duet song while a preceding group was still singing), or in non-aggressive turn-taking duet exchanges (i.e. waiting for the preceding signal to have ended before starting its own one) with neighbours. Under what circumstances, then, do long-distance signallers opt for interference rather than turn-taking? For this purpose, I recorded on a daily basis, gibbons group duet vocal exchanges and investigated the mechanisms that could determine whether a duet song is responded to with turn-taking or counter-singing. If counter-singing occurs, how is it deployed in relation to the previous song? How do groups respond to counter-singing and turn-taking?

My second project was to determine whether or not gibbons could discriminate and thus infer the ‘meaning’ of different contextual song type, duet and predator songs. Nonhuman primates can produce acoustically distinctive vocalisations in different contexts and white-handed gibbons also do so by producing duet and predator songs. Duet and predator songs are identical in the call note repertoire, but consistent differences were found in their overall structure and in the way notes were assembled (Clarke et al., 2006). The main differences lie on song overall duration and the absence of female great calls within the first two minutes of a song in the predatory context. Finer differences were also found in the predominance and production timing of certain note types. As mentioned above, gibbon usually reply by the matching duet song type when hearing neighbouring groups uttering duet song (Raemaekers and Raemaekers, 1985a). Moreover, other observations have also suggested that sometimes gibbons respond to predator songs of neighbouring groups by singing their own predator song, without having spotted the predator themselves (Raemaekers et al., 1984a; Clarke et al., 2006).

To investigate how gibbons encode, discriminate and infer the meaning of structurally different song types by focusing on the production and perceptual discrimination of those two different types of songs, duets and predator songs, I conducted playback experiments. I simulated natural singing events (i.e. duet or predator song) from ‘a simulated neighbouring group’, aiming to elicit in testing groups the matching song types and associated behavioural responses.

Finally, a third project aimed to explore functionally referential alarm calling behaviour in gibbons to refine our actual knowledge about predator song production in gibbons. Alarm calling is a common component of the anti-predator strategies that has evolved in a wide range of species. Alarm calls usually function to announce threats to conspecifics and to communicate directly to the predator (Caro, 2005). In 2009, Manser, defined that functionally referential calls refer to specific stimuli in the external environment of the caller and cause receivers to show an adaptive response to them. Vervet monkeys (*Chlorocebus pygerythrus*) utter acoustically distinct alarm calls to different predator types (Seyfarth et al., 1980a, 1980b), Guereza colobus monkeys (*Colobus guereza*) specify the nature of a perceived danger (eagle vs. leopard) by altering the number of calls per sequence (Schel et al., 2009) and Campbell’s monkeys (*Cercopithecus campbelli*) are able to change the rate of call delivery depending on the level of the perceived danger (e.g. visual vs. auditory predator perception (Lemasson et al., 2010).

Apart from their conspicuous duet songs, the ability of white-handed gibbons to produce song in response to potential predators was identified with the help of different predator models;

large cats, snakes and raptors (Clarke et al., 2006). Because tigers are extremely rare and might have even totally disappeared from Khao Yai National park (Ngoprasert and Gale, 2019), the response to this predator model was excluded from acoustic analysis. While the leopard model elicited strong vocal and anti-predator behavioural responses, the python model never triggered full and loud predator songs, apart from a few utterances composed of soft 'hoo' and 'oo' notes only (3 out of 9 trials), as well as the expected anti-predator behaviour associated (Clarke et al., 2006, 2012). Nonetheless, natural python encounters with gibbons were repeatedly observed. Others and I (Neudenberger, 1993; Uhde and Sommer, 2002) witnessed 9 encounters with real reticulated pythons, which 8 elicited loud calling and 1 with soft calls and strong anti-predation behaviours. Predator model experiments are notoriously susceptible to authenticity problems (Zuberbühler and Wittig, 2011), which could have explained the response differences to model and real encounters with pythons. Despite this potential shortcoming, Clarke et al. (2015) reported subtle differences in song unit arrangements between the clouded leopard and reticulated python songs.

In this study, we followed up on these pilot observations to test more systematically the extent to which white-handed gibbon predator songs are structurally distinct to discriminate between pythons and leopards, the two main predators of wild gibbons. If gibbons have the ability to arrange song units differently to provide reliable referential signals, this would allow others to infer meaning from listening to these songs.

Hylobatids are the first representative to have diverged from the hominoid lineage. Well-studied for decades, they represent a singular subject of interest regarding their complex vocal communication system, make them a particularly relevant model to investigate language-related capacities, such as coding and inference abilities as well as the mechanisms underlying interactive communication between groups.

2- NEPOTISM ENHANCES INTERGROUP COMMUNICATION IN GIBBONS

Manuscript submitted to Animal Behaviour

Julie Andrieu[1], Christof Neumann[1], Suchinda Malaivijitnond[2,3], Ulrich. H. Reichard[4] & Klaus Zuberbühler[1,5]

[1] Department of Comparative Cognition, Institute of Biology, University of Neuchâtel, Switzerland;

[2] National Primate Research Center of Thailand, Chulalongkorn University, Saraburi, Thailand;

[3] Department of Biology, Faculty of Science, Chulalongkorn University, Bangkok, Thailand;

[4] Department of Anthropology and Centre for Ecology, Southern Illinois University, Carbondale, Carbondale, IL, USA;

[5] School of Psychology and Neuroscience, University of St Andrews, UK

* Corresponding author: E-mail: j.andrieu.julie@gmail.com (JA)

2.1 Abstract

Interactive communication is fundamental in all communication systems, and the ability to take turns during acoustic exchanges, a major component of human related-language capacity, has received much more attention in recent years in non-human animals interactive communication. Here, we investigated intergroup interactions in white-handed gibbon duet songs exchanges, which can be competitive or non-aggressive: groups can either overlap another group's ongoing song (counter-singing) or await the end of their song (turn-taking). We found that group composition, relatedness and spatial proximity all played significant roles. Multi-male groups frequently elicited competitive counter-singing from neighbours, whereas turn-taking was frequently elicited by single-male groups, when the recipient group contained genetically related individuals, especially at close proximity. Groups experiencing counter-singing increased their song duration while counter-singing groups did not show any adjustments. We concluded that the default strategy for white-handed gibbons is competitive counter-singing, especially over short distances, but that turn-taking is common when responding to relatives. Perceptions of kinship are powerful cognitive and evolutionary

drivers of human social communication beyond the immediate family units, and our results show the same pattern in a non-human primate, with the first evidence of kinship favouring turn-taking exchanges between groups. Taking turns to communicate might have originated as a kin-selected form of communication.

2.2 Introduction

Kinship has long been recognised as a major driver of cooperation, altruism and division of labour (Lukas and Clutton-Brock, 2018). These complex behavioural phenomena are sometimes also observed between genetically unrelated individuals, for example, if two unrelated mating partners share a genetic interest in a third individual (offspring) or if environmental challenges are better solved at the group level than individually (hunting, antipredator behaviour or territory defence (Fairbanks, 1990; Clutton-Brock et al., 2001; Bergmüller et al., 2007; Townsend et al., 2011). However, kinship is also thought to have played a role in the evolution of language (Fitch, 2005), and the 'information-sharing hypothesis' provides an interesting theory about the way kin selection has evolutionary favoured the complexification of communication capacities (i.e. vocal repertoire expansion) in Hominoids, via vocal learning, by sharing beneficial information with kin (Fitch, 2004; Nowicki and Searcy, 2014).

Indeed, sharing information is not only crucial for many species but it also drives human social interactions. Human dialog mainly consists of exchange of social information, that mostly concerns personal experiences and relationships (Dunbar et al., 1997). Speech exchange requires a minimum of two individual, with a precise-time commitment where avoidance of vocal overlap is crucial for a correct understanding from each part, preventing loss of information (Stokes and Williams, 1968). Conveying and receiving information is decisive in communication and can benefit the receiver which in turn responds to the signaller. This alternation of vocal interactions was termed turn-taking (Sacks et al., 1974) and this framework depict an essential component that was only dedicated to human language due to the lack of specific features as for example grammar and semantic in animal communication. Though bi-directional acoustic exchanges also exist in many species, and the ability to take turns during was recently shown to exist in both, animals and humans, leading to the emergence of numerous comparative studies attempting to depict the evolutionary process that could have driven such capacity in humans.

Turn-taking is fairly common in primate communication, usually when individuals from the same group exchange contact calls (Diana monkeys (*Cercopithecus diana*), Campbell's monkeys (*C. campbelli)* (Uster and Zuberbühler, 2001)(Lemasson et al., 2011), squirrel monkeys (*Saimiri sciureus*) (Symmes and Biben, 1988), and marmosets (*Callithrix jacchus*) (Takahashi et al., 2013; Borjon and Ghazanfar, 2014). Turn-taking has also been reported in gibbons but, unlike in other groups of primates, turn-taking in gibbons not only takes place within (i.e. duetting) (Terleph et al., 2017) but also between groups (Tenaza, 1976; Fan et al., 2006) : Siamangs (*Hylobates syndactylus*) (Geissmann and Orgeldinger, 1997), Mueller's gibbons (*Hylobates muelleri*) (Inoue et al., 2013) , black-crested gibbons (*Nomascus concolor*), Kloss's gibbons (*Hylobates klossi*), Bornean gibbons (*Hylobates muelleri*) (Tenaza, 1976; Mitani, 1985b; Fan et al., 2006); and in white-handed gibbons (*Hylobates lar*) (Raemaekers and Raemaekers, 1985b). Vocal turn-taking in primates and humans is very short, ranging from a few milliseconds to seconds (Ravignani et al., 2019). In gibbons, intergroup duet song exchanges are within several minutes but. regarding the length of their songs, the term 'turn-taking' is still used (Inoue et al., 2013). In humans, turn-taking is a characteristic feature of speech (Ghazanfar and Takahashi, 2014) which can be between individuals with no genetic or social ties. However, unlike other animal communication systems, human speech is semantically much more loaded, suggesting that turn-taking is due to shared interests in extracting the content of each other's utterances (Sacks et al., 1974; Levinson, 2016). Disrespect of turn-taking (i.e. overlapping speech) is also observed in human conversation, presumably in situations of increased social competition (Chowdhury et al., 2015), but this characteristic was also found in many birds species or in the interactive long-distance duet song exchanges between gibbons group, for which song overlap is defined as counter-singing (Tenaza, 1976; Fitzsimmons et al., 2008).

Many species use long-distance calls to interact with outgroup members, and have evolved specialised signal systems for this purpose (McGregor, 2005). Birdsong is a classic example, but other forms of long-distance signalling have also been described. In birds, male song functions to compete with each other over access to females (Catchpole and Slater, 2003) or to defend a territory (Searcy and Beecher, 2009), and males sometimes compete with each other by a process called 'counter-singing' where two males call back and forth to in order to interfere with each other's song (Armstrong, 1973; Greenfield, 1994; Gerhardt and Huber, 2002; Naguib and Mennill, 2010). Although counter-singing makes an individual's own signalling less efficient, it also lowers the efficacy of a competitor's signal by making detection more difficult and possibly less attractive for sexual partners (Ficken et al., 1974; Wilson et al., 2016). In Great tits (*Parus major*), playback experiments of simulated counter-sung opponents showed that individuals increased their signalling effort and costs, by singing longer, compared to alternating playbacks (Dabelsteen et al., 1996). In black-crested gibbons, Kloss's gibbons and Bornean gibbons (*Hylobates muelleri*), songs produced interactively with other groups were longer than songs produce alone (Tenaza, 1976; Mitani, 1985b; Fan et al., 2006;), suggesting that groups are attending and responding to each other's songs. In contrast, when 'turn-taking' callers wait for the previous individual or group to terminate its signalling effort, benefits are in terms of interference avoidance and information loss (Sacks et al., 1974; Yoshida and Okanoya, 2005).

In most gibbon species, the mated pair produces highly synchronised morning songs, so-called duet songs, on an almost daily basis. Duet songs travel over long distances of up to one kilometre (Marshall and Marshall, 1976) and often trigger responses from other groups, which ultimately results in song waves that cover large parts of the forest. Contagion can either be in the form of turn-taking (i.e. waiting to start a duet song until the previous group has finished) or counter-singing (i.e., starting to sing before the previous group has finished). During counter-singing exchanges, groups were often heard to pause for one specific acoustic figure repeated throughout the song bout, the female great call (Raemaekers and Raemaekers, 1985b), though this phenomenon have never been tested. Great calls are individually specific, and were found to be correlated with female physical state, (Terleph et al., 2015, 2016), meaning that groups might gain useful information about outgroup callers. In the wild white-handed gibbons population of Khao Yai National park, outgroup members can be related, with males dispersing from their natal group in neighbouring territories, sometimes as secondary males, whereas females seem to dispersed further, as they have to evict a resident adult female or form a new group (Matsudaira et al., 2018). But until now, it was challenging to investigate whether kinship between outgroup members could influence the modality of duet singing exchange in gibbons. Comparative studies often show considerable interest in our non-human primate closest relatives to assess the evolutionary process by which human 'turn-taking' speech exchange ability might have evolved from predecessors. However, gibbons, family occupying the most basal branch of the hominid lineage, are of great interest to examine duet exchange modalities between unrelated and related outgroup members, taking into consideration the known function of their conspicuous duet songs.

Here, we were interested under what circumstances gibbon pairs opt for turn-taking (instead of competitive counter-singing) and the response strategies of groups targeted by competitive counter-singing. There are two main functional hypotheses for gibbon duet song; territorial defence and sexual rivalry (Raemaekers and Raemaekers, 1985a), both of which are based on the notion of intergroup competition (Mitani, 1985b; Raemaekers and Raemaekers, 1985a, 1985b). In playback experiments, for example, white-handed gibbons responded differently to neighbour duets broadcasted from the border or centre of their territory (Raemaekers and Raemaekers, 1985a). For the territorial function we thus predicted that competitive counter-singing should be negatively correlated with the distance between two calling groups. For the sexual competition hypothesis, we predicted that kinship should decrease counter-singing and favour turn-taking between groups.

2.3 Material and methods

a. Study site and subjects

The study was conducted in the Mo Singto-Klong E-Tau area of Khao Yai National Park, Thailand (101°22'E, 14°26'N), 130 km North-East of Bangkok. Data were collected from December 2012 to August 2014. The local climate is characterised by a wet season from May to October and relatively dry conditions from November to April (Whitington and Treesucon, 1991). Thirteen habituated groups of white-handed gibbons were available for our study, totalling N=53 individuals (see supplementary information, Table S2.1). The population density of white-handed gibbons in the Mo-Singto – Klong E-Tau area is 15.9 individuals per km^2 (Barelli et al., 2007).

b. Population demography

Habituation started in the 1980s, with regular visits of all groups and long-term demographic records of births, migration history and genetic relationships for most individuals (Barelli and Heistermann, 2012; Barelli et al., 2008a, 2008b, 2013b; Brockelman et al., 1998; Matsudaira et al., 2018; Reichard, 1995, 1998, 2009; Reichard and Barelli, 2008; Reichard and Sommer, 1997; Reichard et al., 2012; Savini et al., 2008; Uhde and Sommer, 2002).

Individual's genotypes were determined in a previous study by the extraction of three genetic markers (mitochondrial DNA, Y-chromosomal haplotypes and autosomal microsatellite), which were subject to Mantel tests to identify kinship between individuals of 17 groups

(Matsudaira et al., 2018). Since 2011, when genetic samples were collected, all migration events and changes in group composition were documented (see supplementary information, Figure S2.1).

c. Terminology

A duet song is defined as vocal behaviour uttered conjointly and in synchrony by the mated pair. Duet songs are given mostly mid-morning, peaking around 08:30 local time, anywhere in the territory (Raemaekers et al., 1984a), usually lasting about 10 min (Clarke et al., 2006). Each duet is a structurally complex utterance divisible into (i) an introductory sequence (lasting about 2 min, containing most of the pair's song notes, starting with soft "hoos" and ending with the first female great call (Raemaekers et al., 1984a; Clarke et al., 2006), (ii) the female's great call sequence (lasting from 12 to 22 seconds, containing from 6 to 13 notes, usually followed by the male's response coda), and (iii) the interlude sequence (containing mostly 'wa' notes phrases separated by pauses). Duet songs can then contain more great call/interlude sequences and end with either of them (Ellefson, 1968; Marshall and Marshall, 1976; Raemaekers et al., 1984a).

A duet can be produced spontaneously or in response to an ongoing or previous duet by another group, oftentimes until large parts of the forest are covered by singing gibbon groups ('duet waves'). Response duets can be initiated while the previous group is still singing (counter-singing) or after a group has stopped singing (turn-taking) between 10s (i.e. duration of the introductory 'hoo' note series duration, barely humanly audible at distance (Andrieu et al., 2020) and 5min (i.e. definition of the maximum delay threshold of a response (Raemaekers and Raemaekers, 1985b)) after the completion of another duet (Figure 2.1). If more than two groups produced their songs, either by overlapping or turn-taking, we considered the songs to be part of a duet wave. If there was a period of silence exceeding 10min, we considered a duet wave to be completed, by doubling the response threshold time, we ensure that the next song is the initial song of the wave.

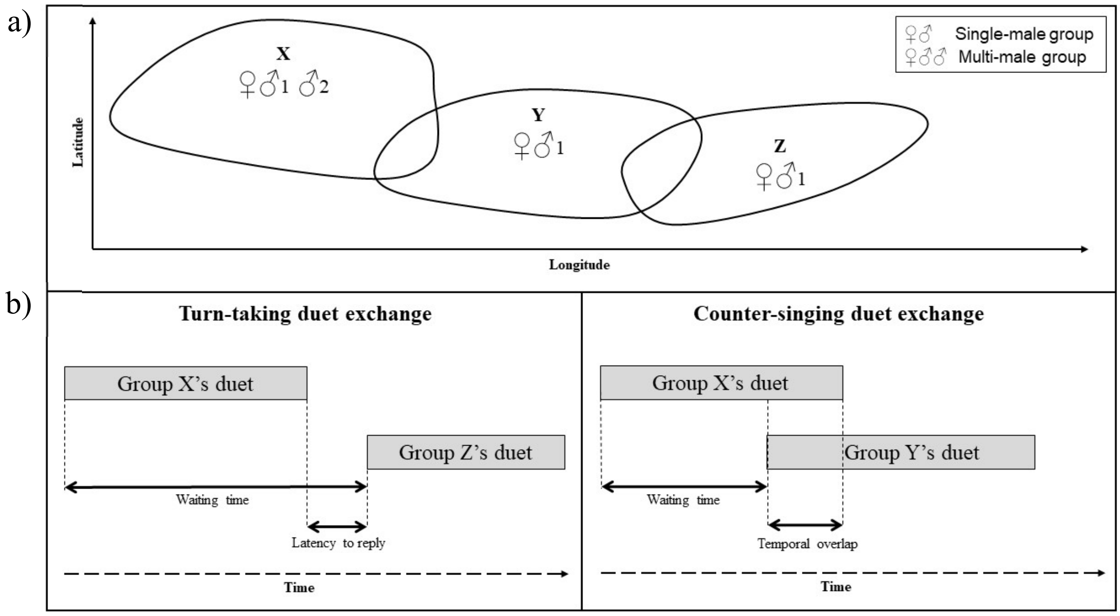

Figure 2.1. Schematic illustration of (a) spatial relations, group composition and (b) turn-taking and counter-singing between gibbon groups (Groups X, Y and Z; 1= primary male, 2 = secondary male).

d. Data collection

Territory mapping

The thirteen study groups were regularly followed from December 2012 to August 2014. Whenever a group was found (N=217 contacts over 263 days), we recorded its position at 5-min intervals to determine the group's home range and centroid. We used a hand-held Global Positioning System receiver (GPSmap®62s, Garmin International, Inc, Olathe, Kansas City, Missouri, U.S.A) to determine the Universal Transverse Mercator (UTM) units. We used the Adehabitat HR package (Calenge, 2011) in **R** (V3.5.1 (R core Team, 2018) to analyse the GPS data. As home range estimators, we used Minimum Convex Polygons (MCP) and 95% Kernel Density Estimation (KDE) to determine territory boundaries and areas (see supplementary, Figure S2.1).

Acoustic data

We recorded song events on an all-occurrence basis (Altmann and Altmann, 1977), using directional microphones (Sennheiser MKH815T & Sennheiser ME66) with windshields connected to a digital stereo recorder (Marantz PMD660) at a 44.1 kHz (16 bits) rate, from December 2012 to August 2014. Whenever a song was produced, we noted the identity of callers and the GPS position of the focal group, as well as the identity of any neighbouring group heard previously emitting song or responding to the focal group's song, including their estimated distances relative to the focal group. For each song, we determined (i) the song type (duet song, ooaa duet, male solo song, female solo song or predator song) (Clarke et al., 2006), (Raemaekers et al., 1984a), (Raemaekers and Raemaekers, 1984b); (ii) the onset and offset time of the song, (iii) the number of great calls produced in each song; (iv) relevant external events, such as predation attempts or group encounters.

We recorded songs over N=194 days (corresponding to 1,007 hours of recording). From this dataset, we excluded low-quality recordings, that is, recordings made during strong winds (N=9 days) or heavy rain (N=11 days) or while experiencing technical problems (N=3 days). Recordings were analysed with Raven Pro 64 1.4 (Cornell laboratory of Ornithology, Ithaca, NY) and cross-checked with field notes for (i) group identity (great calls are acoustically unique for each female and can be used as identity markers (Terleph et al., 2015), (ii) onset of each song (+/- 10 s), determined by the occurrence of the first audible "hoo" notes series. Hoos are relatively quiet notes of low frequency, which are audible to a human only at close distance, usually when standing under a focal group, (iii) song offset and corresponding song duration; (iv) inter-song intervals as the duration in seconds between the offset of a song and the onset of a subsequent song, (v) song overlap as the amount of time (s) between the offset of the previous and onset of the next song. Spectrograms were set for fine temporal resolution (Hamming window function; 256-point fast Fourier transformation).

e. Data analysis

We restricted analyses to the first two duet songs (initiator and response) given during each duet wave. Songs that did not qualify as duets (predator songs, adult male and female solo songs, ooaa duets) were excluded from the analysis (Clarke et al., 2006), (Raemaekers et al., 1984a), (Raemaekers and Raemaekers, 1984b). To distinguish duets given as part of turn-taking or counter-singing, we compared the first two duets in each singing wave and assigned to each: a) group identity, b) group composition (i.e., single vs. multi-male group), c) distance between territory centroids, d) genetic relatedness between the recipient male and initiator male and female, e) onset and offset time (s) of each duet, g) interval between the offset of the initiator's duet and onset of the response duet, and h) and weather condition.

During the study period, we did not observe groups with more than two adult males, except for the rare case when in addition to two adult males a mature male offspring ready to disperse was present in the group. Mated females were unrelated to either the primary or secondary males of their own group. Among the 13 study groups, we found 3 female-female pairs who were related to each other, but this sample was too small to include female relatedness into systematic analyses of the effect of kinship. In contrast, we observed 9 male-male relatives and 7 male-female relatives, perhaps because of males' propensity to disperse more locally than females (Matsudaira et al., 2018). For both male-male (M-M) and female-male (F-M) intergroup song interactions, we considered genetic relatedness from the responding male's perspective (M-M, n= 9 group pairs), (F-M, n=7 group pairs) (see supplementary, Figure S2.1).

We calculated minimum and maximum distances between borders and core areas, for neighbouring and non-neighbouring groups (see supplementary, Figure S2.1). We added weather condition to our dataset, assessed every two hours as either 'good' (sunny, no wind) or 'poor' (rainy and/or windy). This was because gibbon singing appeared to be affected by weather, with less or no singing during poor compared to good weather (Brockelman and Srikosamatara, 1993), possibly due to differences in sound propagation (Brunet et al., 1996).

Finally, to test whether or not counter-singing was affected by female great calls, we recorded 45 duets of focal groups during counter-singing exchanges, either in the role of initiator or responding group. We collected, for all great call sequences produced by the focal group's female, a) the identity and group composition of the focal group; b) whether the focal group was initiating or responding to counter-singing; c) the identity and composition of the counter-singing group, and d) whether the counter-singing group was a neighbour or not a neighbour. We assigned an ordinal number to each great call produced by the female of the focal group and noted whether the counter-singing group paused for the great call sequence of the focal group's female.

Statistical procedures

Statistical analyses were performed using **R** (V3.5.1 (R core Team, 2018)) with the lme4 package (Bates et al., 2014). All the statistical models used are described below, and for further description, see supplementary (Table S2.2).

Model 1: What determines cooperative versus competitive singing?

To investigate factors determining counter-singing versus turn-taking, we fitted generalised linear mixed models (GLMMs) with binomial error structure and logit link function (response variable: probability of counter-singing, i.e., whether a group started singing before the initiator group had stopped singing). As predictor variables we used: 1) the distance between home range centroids, 2) the composition of both groups (two-level factor), 3) the genetic relatedness of males of both groups, and 4) the genetic relatedness between the initiator female and the response male. 5) We included a two-way interaction between group compositions, reasoning a difference in male group composition (single vs. multi-male groups) might influence male intragroup sexual competition in multi-male groups. Furthermore, we incorporated interactions between group distance and the two genetic relatedness variables between the responder male and both, initiator male and female. Finally, we fitted a 3-way interaction between distance between group centroids and group compositions (single-male versus two-male group). We included weather condition as a control variable (two-level factor); signaller and recipient group identities were used as random intercepts. We fitted the following uncorrelated random slopes in recipient group identity: distance between groups (z-standardised), composition of the initiator group and the two relatedness variables. Residuals of the full model were visually inspected for normality or homoscedasticity and no obvious deviation from normality was detected. We compared this full model to a null model with the same random structure but only weather condition retained as fixed effect. If comparisons revealed a significant difference, we tested and removed non-significant interaction terms. Finally, we applied a leave-one-out procedure where we refitted the model multiple times, with data from one group at a time removed, and then compared these results to those of the full model. We then applied a leave-one-out procedure to investigate whether any of the groups singing in second position affected disproportionally the results of the full model.

Model 2: What determines the onset of response duets?

How long did the responder group wait until it started with its own duet in relation to group composition, relatedness and distance between territories? To explore the temporal sequence of subsequent duets, we fitted a LMM investigating the duration between the onsets of initiating and response duets. We squared transformed durations to achieve an approximately normal distribution before fitting a linear mixed model (LMM). As predictor variables, we used distance between group territories, composition of both groups (two-level factor), relatedness variables and weather as a control variable. As in *Model 1*, we incorporated a 3-way interaction between distance and both groups' composition and a 2-way interaction of distance and relatedness. Before fitting, we centred and standardised distance from centroids for both groups. We included both group identities as random intercepts. We fitted uncorrelated random slopes in responding group identity: distance between groups (z-standardised), composition of the initiator group and the two relatedness variables. We visually inspected residuals of the full model for normality or homoscedasticity and detected no obvious deviation from normality and fitted a null model that we then compared to our full model with a likelihood ratio test (LRT) and we removed non-significant interaction terms.

Model 3: Do counter-singing groups pay attention to initiator songs?

To test whether counter-singing groups pay attention to the female great call sequence, we fitted a GLMM with binomial error structure and logit link function. The response variable was whether the counter-singing partner paused during the focal group's great call sequence (i.e. the focal group can either be in the role of initiator (scenario 1) or responder group (scenario 2). As predictor variables, we used the focal group's role (scenario 1 and scenario 2), neighbour vs. non-neighbour, group type composition of both groups (single vs. multi-male) and focal group's great call production position order in the duet. We included the following two way-interaction terms: scenario x neighbourhood status; scenario x group composition of both groups; and scenario x target group's great call production order. Group identities were used as random intercepts and weather as a control variable. We additionally fitted the following uncorrelated random slopes in target group identity: singing order of the target group (scenario), group composition of the counter-singing partner, great call position order and the neighbourhood status; and in counter-singing group identity: singing order of the target group (scenario), group composition of the target group, great call position order and the neighbourhood status. We fitted a null model that we compared to our full model with a likelihood ratio test (LRT). If the test revealed significance, we then tested interaction terms and removed non-significant interaction terms. We fitted a null model that we compared to the full model and checked assumptions about normality or homoscedasticity and detected no obvious deviation from expected normality.

Model 4: How do initiator groups respond to counter-singing?

To test whether initiator duet groups experiencing counter-singing duetting responded with counter-measures, we classified initiator duets into three categories: (a) isolated duet (i.e. no duet response: control condition), (b) duet followed by turn-taking duet (i.e., subsequent duet starts after initial duet had ended), and (c) duet followed by counter-singing duet (i.e. subsequent duet temporally overlapping with initiator duet). We then fitted a linear mixed model (LMM) to analyse variation in initiating duet duration, that we squared transformed to achieve approximately normal distribution. Predictor variables were the duet category (three-level factor) and the initiator group composition and their interaction. As the models before, we included weather condition as a control variable, included random intercepts for group identity and fitted uncorrelated random slopes for duet category in group identity. As the null model, we retained the identical random structure and included only weather condition as fixed effect. As in the models before, we then compared the full and null models with a likelihood ratio test (LRT). If this comparison revealed a significant difference, we continued by testing the interaction term with a LRT and removed non-significant interactions. We tested remaining terms with LRTs and visually inspected residuals from the full model to check assumptions about normality or homoscedasticity for which no obvious deviation was detected. Finally, we applied a leave-one-out procedure where we refitted the model multiple times, with data from one group at a time removed, and then compared results with those of the full model.

Model 5: How do counter-singing groups respond to initiator group responses?

We investigated whether or not counter-singing groups adjusted song duration compared to when they were turn-taking. We fitted a LMM with the duration of the responder duet as the response variable (squared transformed). We distinguished between cases were the responder group (a) was not followed by a subsequent third group's duet (i.e., no further duet wave, control condition), and (b) was followed by a subsequent third group's duet but with no overlap. As predictors, we used whether or not the responder group interrupted the initiator duet (i.e., the interruption variable: counter-singing or turn-taking), group composition of both groups, the duet category of the responding group (i.e. responder duet with no further wave, and responder duet with a third subsequent duet with no overlap, to be sure that responder duet won't be affected by the interruption of third subsequent duets) and weather condition (control factor). We also fitted two two-way interactions: responding group's duet category with responding group composition and responding group's duet category with the interruption variable. Group identities of initiator and responding groups were used as random intercepts. In addition, we fitted the following uncorrelated random slope for responding group: initiator group composition, the interruption variable and responding group duet category. We fitted a null model, compared to the full model and removed non-significant interaction terms as described above.

2.4 Results

Duet songs were more commonly initiated by males than females (71.9% vs. 28.1%, respectively; N=32 duets randomly chosen in our data set) regularly resulting in duet waves in 61.7% of cases (vs. 38.3% for isolated duets, N=410). Overall, counter-singing was about twice as common as turn-taking (63.2% vs. 36.8%, respectively; N=253), especially between neighbours compared to non-neighbours (80.0% neighbours vs. 20.0% non-neighbours, respectively, N=135, see Table S2.3 for between group distances).

a. Model 1. What determines cooperative versus competitive singing?

We tested whether the probability of groups to interrupt an initiator duet was affected by the distance between the groups' 'territories centroids, the number of males per group, their relatedness and the relatedness between the initiator female and responder male. The full model differed significantly from its null model (GLMM: $\chi^2(11)=37.61$, $P<0.001$). We found a significant 3-way interaction between (standardised) intergroup distance, composition of the initiator group and composition of the responder group (GLMM with LRT: $\chi^2(1)=9.01$, $P=0.003$, Figure 2.2).

Counter-singing duetting was more common with smaller than larger distances, but only if at least one group was single-male. When both groups were multi-male (n=20), counter-singing duetting was generally common, but more so at large than small distances (Figure 2.2). Note that the distance variable refers to standardised, not real, distances to allow comparisons across conditions.

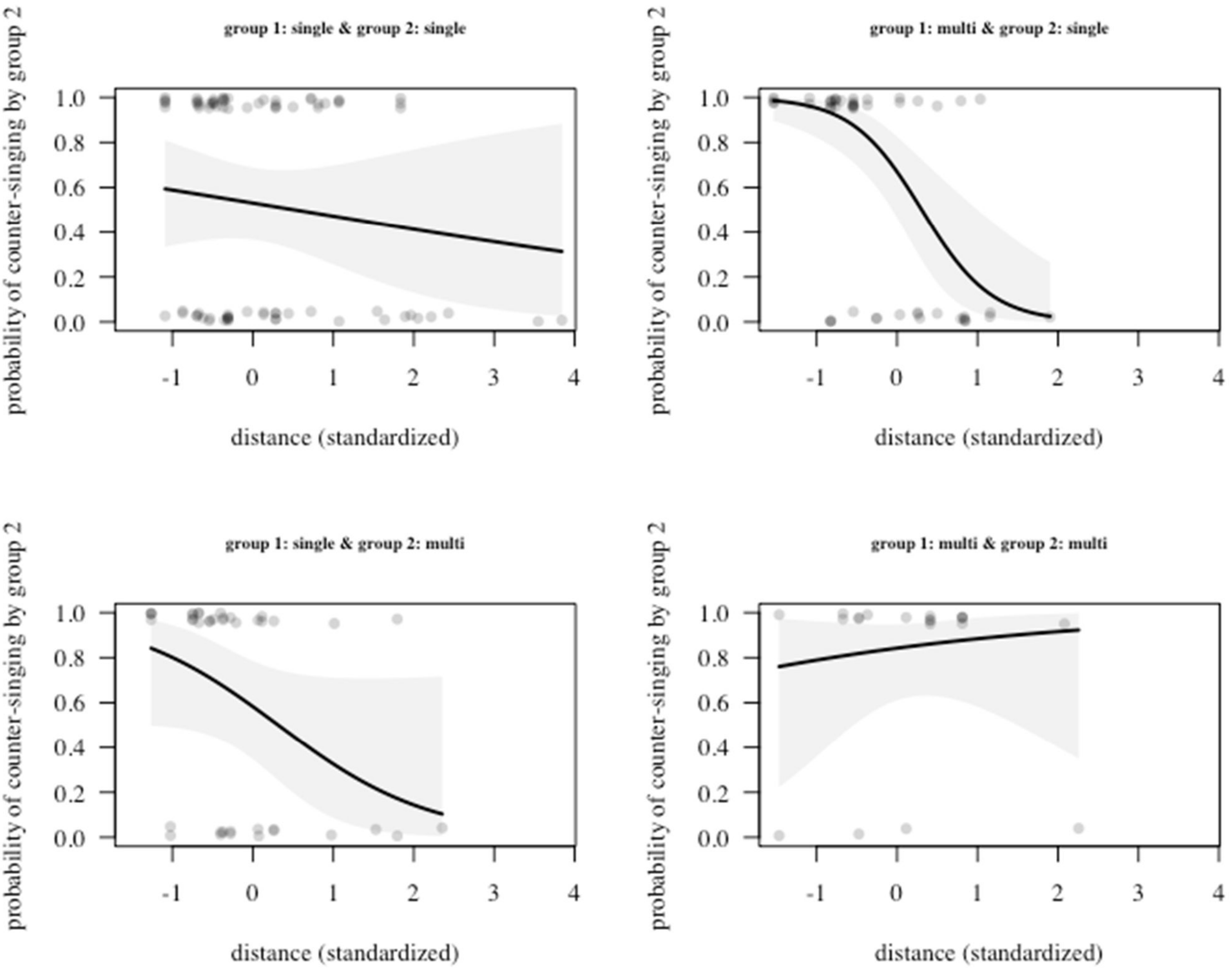

Figure 2.2. Effect of three-way interaction between intergroup distance and group composition of both, initiator and responder groups (group 1= initiator; group 2= responder) on the probability of counter-singing to occur (single= single male group; multi = two-male group). Lines are fitted model estimates and shaded areas are 95% confidence bands.

Second, we found a significant two-way interaction between intergroup distance and genetic relatedness between males of the different groups (GLMM with LRT: $\chi^2(1)=6.81$, $P=0.009$, Figure 2.3). If males of neighbouring groups were kin, counter-singing duetting was less frequent at small compared to large distances (Figure 2.3 left panel). In contrast, if the two males were unrelated, counter-singing was more common at close compared to large intergroup distances (Figure 2.3 right panel).

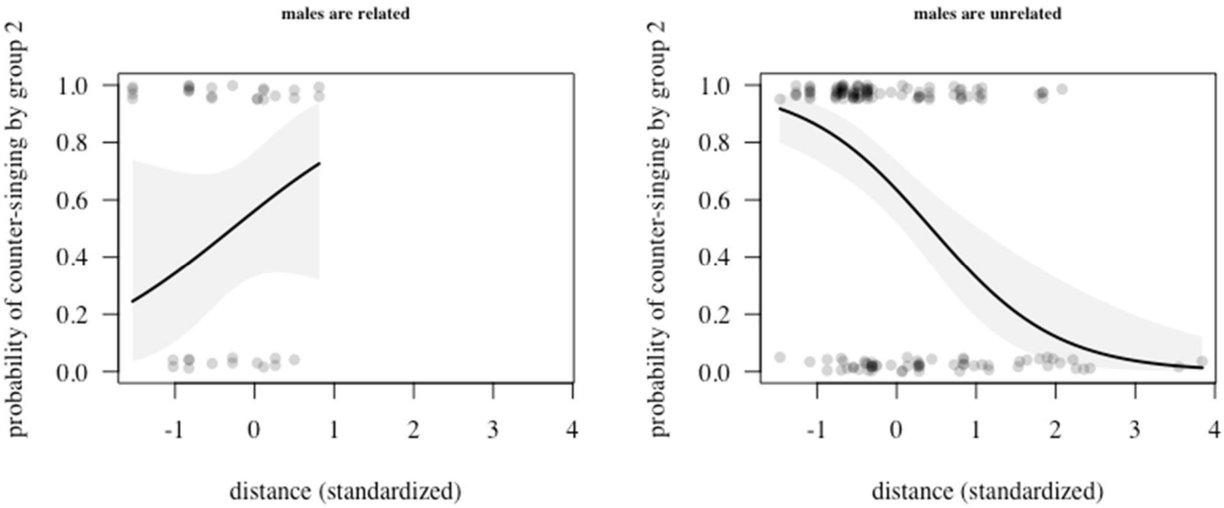

Figure 2.3. Interaction between intergroup distance and genetic relatedness between initiator and responder males on the probability of responder groups to engage in counter-singing exchanges. Lines are fitted model estimates and shaded areas are 95% confidence bands.

Third, we found an effect of relatedness between initiator group's female and the responder group's male, with unrelated pairs being more likely to counter-sing duets than related pairs (GLMM with LRT: $\chi^2(1)=5.40$, $P=0.020$, Figure 2.4).

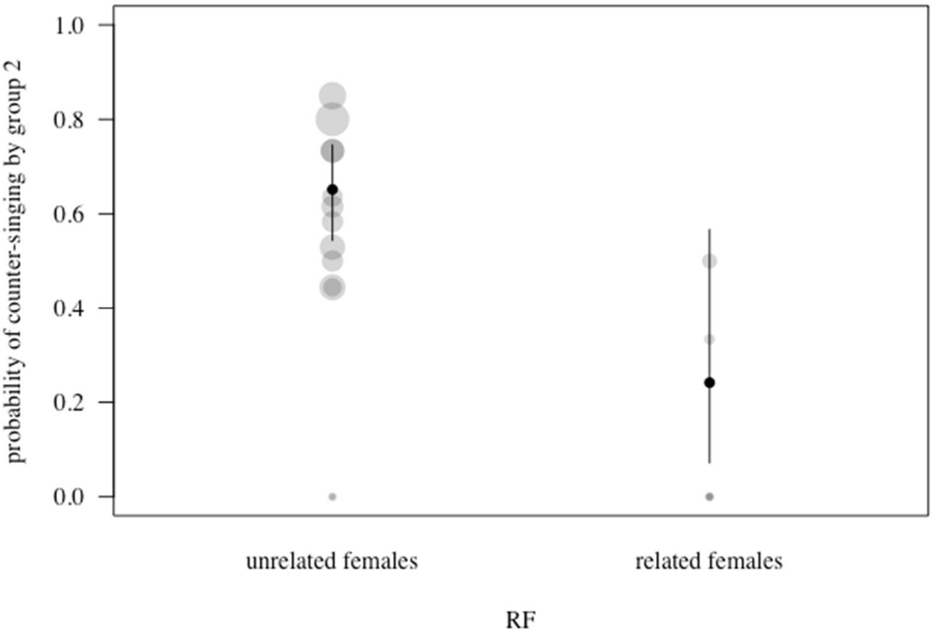

Figure 2.4. Effect of kinship between initiator females and responder males on the probability of responder groups to engage in counter-singing exchanges.

b. Model 2. What determines the onset of response duets?

In duet exchanges, we examined how groups determined their duet onset with respect to the initiator group's duetting. The full model (variation in waiting time) differed significantly from its null model (LMM: $\chi^2(11)=22.76$, $P=0.019$, Figure 2.5). When non-significant interaction terms ($\chi^2(6)= 7.53$, $P = 0.275$) were removed, we found that the number of males in initiator groups affected the duet onset time of responder groups' singing (LMM with LRT: $\chi^2(1)=4.01$, $P=0.045$, Figure 2.5 right panel), with a longer wait time if the initiator group was a two male group. In addition, larger distances between groups (LMM with LRT: $\chi^2(1)=6.47$, $P=0.011$, Figure 2.5 left panel) affected responder groups' singing. In contrast, the number of males in responding groups had no effect (LMM with LRT: $\chi^2(1)=2.31$, $P=0.129$) and neither did relatedness (LMM with LRT, between males: $\chi^2(1)=0.05$, $P=0.829$ or between initiator duet group's female and responder group's male: $\chi^2(1)=2.10$, $P=0.147$).

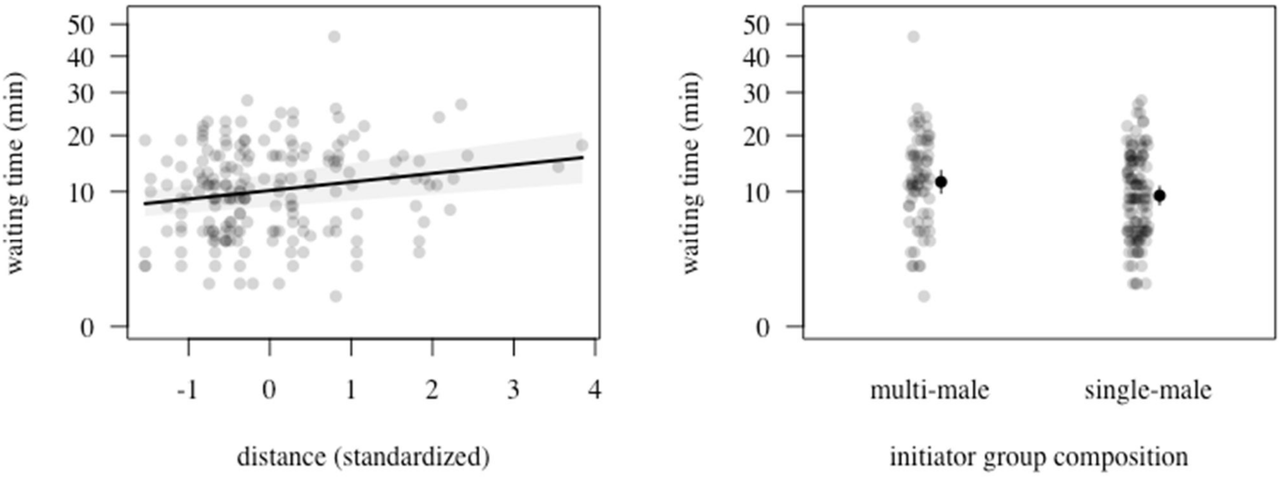

Figure 2.5. Effect of distance and initiator group composition on responder groups waiting time duration.

c. Model 3. Do counter-singing groups pay attention to initiator songs?

Song duets in white-handed gibbons contain a characteristic vocal feature produced by the female, the great call sequence. Great calls are usually given within the first 2 minutes of duet singing (Clarke et al., 2006; Raemaekers et al., 1984a). They are individually distinct, allowing listeners to infer a singing female's identity (Terleph et al., 2015). We examined N=45 duets produced by focal groups during counter-singing duet exchanges, either in the role of initiator duet group (scenario 1) and responder group (scenario 2); and investigated whether or not counter-singing groups paused upon focal group female's great calls (N=303 great call sequences).One hundred and ninety great call sequences were partially overlapped by the other group's singing (scenario 1: 51.1%; scenario 2: 48.9%). When comparing the full model (pausing for great calls) to its null model we found a near-significant difference (GLMM: $\chi^2(9)=16.61$, $P=0.055$). We therefore removed the non-significant interaction terms and found that pausing for great calls was most common between neighbouring compared to non-neighbouring groups (GLMM with LRT: $\chi^2(1)=5.99$, $P=0.014$, Figure 2.6 left panel).

Finally, we found a trend concerning the two-way interaction between the focal group singing scenario and their great call position order in the duet (GLMM with LRT: $\chi^2(1)=3.52$, $P=0.061$, Figure 2.6 right panel). It appears that when focal groups are initiator duet groups in the counter-singing dyad (scenario 1), responder groups had a greater probability to pause for initiator female's great calls whatever the position of great call throughout the duet. On the other hand, however, the probability of initiator groups to pause on focal female's great call sequences when focal groups are singing in second position (scenario 2), appears to decrease as the duet bout progressed.

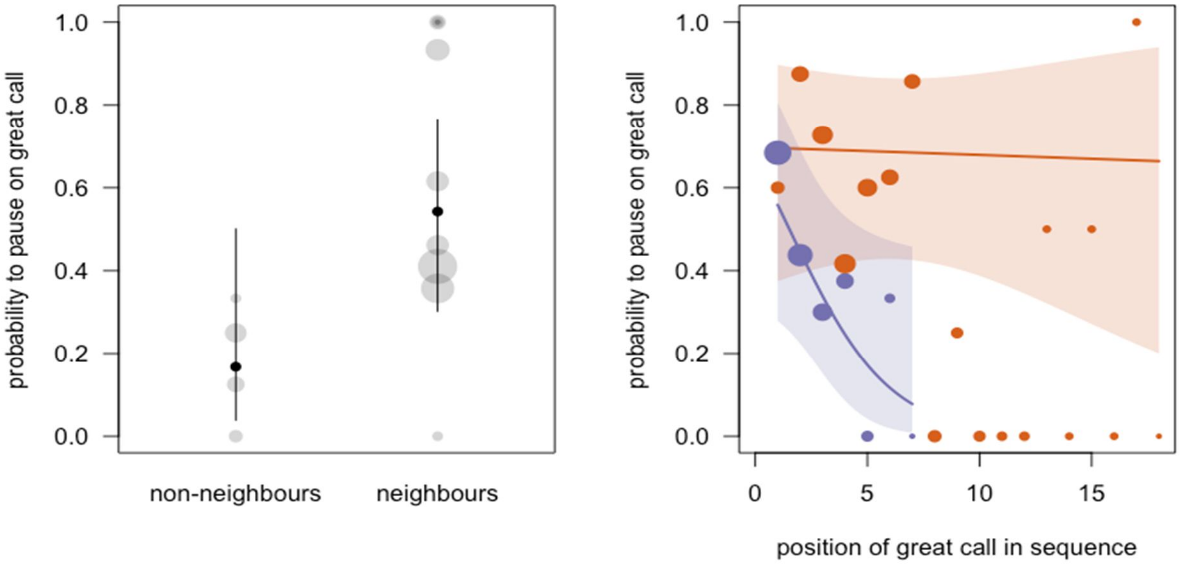

Figure 2.6. Neighbourhood effect on counter-singing duet groups and effect of the two-way interaction between focal group singing order. Red dots and line represent scenario 1 = focal group singing first, violet dot and line represent scenario 2 = focal group singing second) and their great call position order produced throughout duets, on the probability of groups to pause for great call sequences of the focal group.

One should keep in mind that the latter result has to be taken very cautiously as we only found a trend concerning the two-way interactions. While only closely looking at the first scenario, it seems that the probability of responder groups to pause on focal group's great calls when focal groups sing first also seem to decrease as the bout progress (see supplementary, Figure S2.2). Nevertheless, and for the sake of having good recordings quality (with one observer being under one of both group in the counter-singing dyad), we could not analyse the first two duet per wave as in other models, and other duets could have interfered with the probability of groups to pause on great call sequences. However, these results are still promising for future analysis with a more appropriate design.

d. Model 4. How do initiator groups respond to counter-singing?

To test whether initiator groups adjusted their singing when experiencing counter-singing duets, we classified initiator groups duets (N=389) as isolated duets (n=157, control condition), turn-taking duets (n=93) and counter-singing duets (n=139). We then tested whether initiator duet duration was affected by the singing behaviour of a subsequent group (no response, turn-taking and counter-singing) in interaction with initiator group composition. Our full model (initial duet duration) was significantly different from its null model (LMM: $\chi^2(5)=39.49$, $P<0.001$), suggesting that the interaction of both factors might affect initiator duet groups' singing duration. We therefore tested the significance of the interaction term (LMM: $\chi^2(2)=4.11$, $P=0.128$), and found that the pattern of response of the responding group and the initiator duet group composition distinctly affect the initiator group's duet duration (Figure 2.7). First, we found that overlapped duets were about 5 minutes longer than both isolated duets and non-overlapped duets (LMM with LRT: $\chi^2(2)=32.74$, $P<0.001$, Figure 2.7 left panel). Then, and independently of the type of exchanges between initiator and responder groups (i.e. counter-singing or turn-taking), we found that duets produced by multi-male groups were generally about two minutes longer than those uttered by single-male groups (LMM with LRT: $\chi^2(1)=0.019$, $P<0.05$, Figure 2.7 right panel).

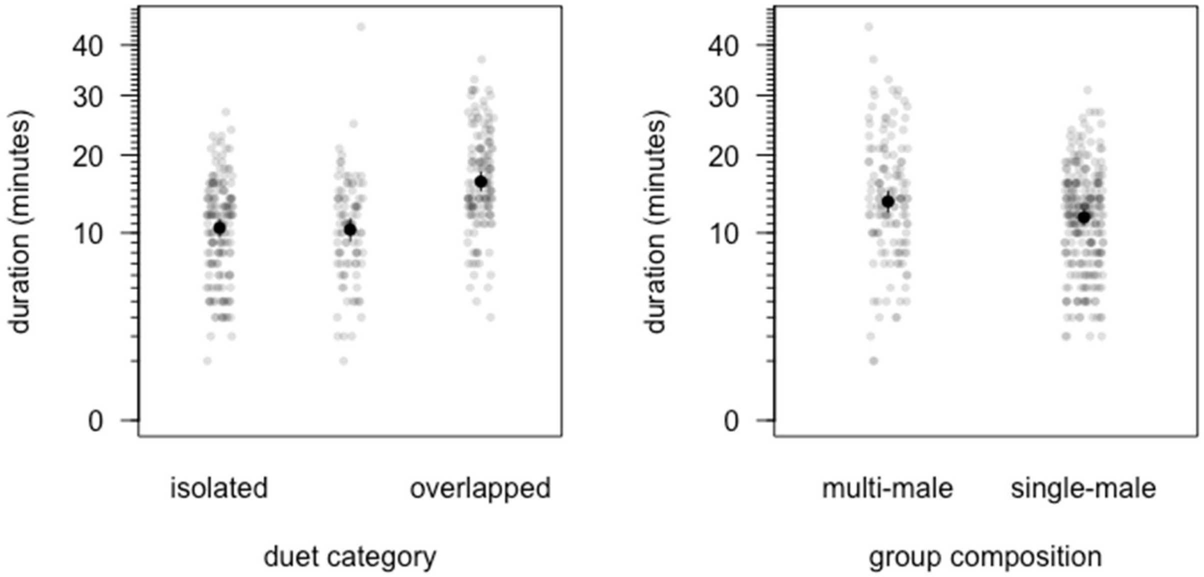

Figure 2.7. Effect of the responding group duet response pattern and initiator group composition on duet duration of initiator duet groups' singing.

e. Model 5. How do counter-singing groups respond to initiator group responses?

To examine whether responder groups adjusted their songs, depending on whether they engaged in turn-taking or counter-singing, we compared their duet durations during both types of interactions. The full model did not significantly differ from its null model regarding duet duration (LMM: $\chi^2(6)=4.77$, $P=0.573$), suggesting that responder groups' songs were unaffected by cooperative or competitive interactions.

2.5 Discussion

In this study, we investigated long-range vocal interactions between groups of white-handed gibbons at Khao Yai National Park, Thailand. Intergroup vocal interactions consist of sequential, fine-tuned duet singing of the pair-bonded males and females, which can be either non-aggressive (i.e. turn-taking) or competitive (i.e. counter-singing) towards neighbouring groups. We found that counter-singing duetting was the default while turn-taking only occurred in some between group interactions. We found that distance between the two groups had an effect insofar as neighbours and spatially close non-neighbours were highly likely to engage in counter-singing. As distance increased, the likelihood of counter-singing decreased, an effect moderated by group composition (single vs two-male groups). In particular, the two groups differed in terms of male composition, the likelihood of counter-singing was negatively related to intergroup distance. The same pattern was also found between single male groups, but not between multi-male groups where counter-singing was common regardless of distance. Moreover, we found that kinship affected the likelihood of competitive counter-singing in two ways. First, if males were unrelated, responder groups were likely to engage in competitive counter-singing, especially at close distances. However, if males were related, we found the opposite effect: responder groups engaged in turn-taking singing exchanges at close distances and counter-singing when distance increases. Kinship between the initiator females and the male responders also increased turn-taking, but this time regardless of distance.

We found that counter-singing in gibbons is not a mere chance event, as suggested for birds (Searcy and Beecher, 2011), but our findings are more in line with Naguib and Mennill (2010) birds theory, meaning counter-singing could serve as a behavioural strategy in intergroup competition. The fact that groups with mutually unrelated individuals engage in counter-singing at close distance is in line with earlier reports that have interpreted gibbon duets as a ritualised form of intergroup spacing (Mitani, 1985b), a type of aggressive signalling that substitutes physical encounters (which were rare after counter-singing: 6.1%; N=49 encounters). Our data thus support the aggressive signalling hypothesis which has been proposed for counter-calling in other animal species (Mennill and Ratcliffe, 2004a, 2004b; Kunc et al., 2006; Naguib and Mennill, 2010). This also explains the fact that individuals targeted by counter-calling respond by increasing song duration (Dabelsteen et al., 1996; Hall et al., 2006) or approaching the opponent (Dabelsteen et al., 1997; McGregor et al., 2000; Mennill and Ratcliffe, 2004a), an effect also found in this and other gibbon studies (Tenaza, 1976; Mitani, 1985b; Fan et al., 2006). One source of competition is caused by substantial home range overlap in this population (0.12-4.66 ha, see Figure S2.1), presumably due to shared feeding sites. We did not test the predictions of this hypothesis in more detail, e.g. whether counter-singing was more common when food was available in contested overlapping areas compared to other times.

A second source of competition is in terms of reproductive opportunities. Here, our data show that multi-male groups were more frequently targeted by counter-singing than single male groups, in line with the mate competition hypothesis. During group encounters males commonly chased each other, while females and their offspring stayed away (Reichard and

Sommer, 1997). Multi-male groups may therefore be perceived as especially threatening because it may be harder for single males to prevent two males from interacting with their female partners, especially because one of the two males is usually only weakly bonded with the female of this group, thus increasing the risk of extra-pair copulations during intergroup encounters (Reichard, 1995; Kenyon et al., 2011; Barelli et al., 2013b). Single males have higher androgen level than males living in multi-male groups (Barelli and Heistermann, 2012). Since androgen levels affect the acoustic structure of songs and other vocalisations (Reby et al., 1998; Fischer et al., 2004; Barelli et al., 2013a) , it is possible that recipients can discriminate single from multi-male groups without visual encounters. Also, multi-male groups duet longer than single male groups, which increases the probability of (arbitrary) interruptions due to counter-singing. Nonetheless, single male groups should gain advantages to deploy counter-singing as a way to prevent future agonistic encounters, and above all with multi-male groups. The mate competition hypothesis, however, is at odds with the finding that relatedness between the initiator and responder group decreased competitive counter-calling. One possible explanation is that males refrained from counter-singing if the neighbouring female was a mother, sister or half-sister, which may have rendered here irrelevant as a mating partner (Gil and Gahr, 2002). Related to this, in competitive counter-singing, groups generally tend to pause for the females' great call, especially during the beginning of a song, suggesting that listeners sought to extract the identity of the calling female (Terleph et al., 2015) her menstrual cycle state and overall physical condition (Terleph et al., 2016) or variables related to the bond with her existing partner (Haraway and Maples, 1998).

Our most striking finding was the effect of kinship between adult individuals on vocal behaviour, largely in two ways. First, turn-taking was favoured when males of both groups were kin, it appears that the responder males are more tolerant to their kin neighbours at close distance; and as distance increase, their probability to overlap kin increases. Counter-singing seems not to be beneficial for genetically related groups at close distance, first because male kin might be not considered as a territorial threat as turn-taking is favoured at close distance; and in the mate attraction hypothesis, as potential extrapair copulation cases exist in gibbons (Barelli et al., 2013b), male gene dispersion would still not be as costly as if it were non-kin male. As distance increases, the probability of counter-singing further kin groups increases. So far, one potential explanation could be that counter-singing might have a function in maintaining contact with non-neighbours kin. White-handed gibbon males were found to have acoustic inter-individual differences (Barelli et al., 2013a), and as male related individuals remain in contact for years in their natal groups before dispersing (father-son: 8-10 years, siblings or half siblings: 2-8 years in our population) and that paternal discrimination was found to be relevant in father/son and also siblings and half siblings combinations, with a 'high probability to interact with each other's' (Widdig, 2007); there is a probability that male related groups might use counter-singing at further distance to maintain contact. Interacting with further non-neighbouring related groups could also be a mean to groom at distance, a mechanism similar to what was found in Japanese macaques (*Macaca fuscata*), where distance vocal exchanges were found to be correlated with social cohesion (Arlet et al., 2015).

Kinship between responder males and initiator group females was also found to favour turn-takin duet exchanges but this time, regardless distance. As mentioned earlier, responder groups might as well avoid overlapping female kin (mother/son and brother/sister in our population) for attracting a potential mate, preventing inbreeding and costliest effort to kin into vocalising longer when overlapped by a subsequent group. So far, our study investigates the kinship effect from the responding males perspective in a in between group framework due to the small sample size of related females in our population (short dispersal distance of males compared to females (Matsudaira et al., 2018), but we agree that the latter point is understudied and needs further investigation. Nevertheless, to our knowledge, this is the first study that explore the role of genetic in a between groups interactive communication system, pointing out turn-taking abilities as a kin-selected form of communication in a non-human primate.

Hylobatids provide an interesting model for the evolution of human communication, due to similarities in social systems, including strong pair-bonds, pair-living structure, bisexual dispersal position in the hominoid lineage. In sum, our study demonstrates that white-handed gibbons at Khao Yai can overcome a basic competitive attitude towards neighbouring individuals and groups by engaging in turn-taking duetting with their long-distance duet songs. Importantly, neither the resource competition nor the mate competition hypotheses can explain fully why related individuals should refrain from competing with each other over access to mating partners or food. Kin selection, in other words, have been partially responsible for the evolution of gibbon song. Comparative studies investigating turn-taking capacities in relation to the origin of human language is a quite limited field; methodological approaches are not standardized, and thus difficult to compare (Pika et al., 2018; Ravignani et al., 2019), and standardized experimental protocols are needed. Whether this capacity could be a shared trait with hominid predecessors or the result of convergent evolution, which latter theory seems to be favoured (Wacewicz and Żywiczyński, 2018), kin selection might have been a characteristic feature that could have led to the emergence of more complex vocal repertoires (Fitch, 2004; Nowicki and Searcy, 2014) but also turn-taking abilities in Hominoids.

2.6 Acknowledgments

We are grateful to the National Research Council of Thailand (NRCT), the National Park, Wildlife and Plant Conservation Department (DNP), and the superintendent of Khao Yai National Park for giving us the permission to conduct this research. We thank Kazunari Mastudarai for his support in the field and Claudia Stephan and Noemie Lamon for their comments and support. The research has been funded by the European Research Council (grant number FP7; PRILANG GA283871).

2.7 Supplementary information

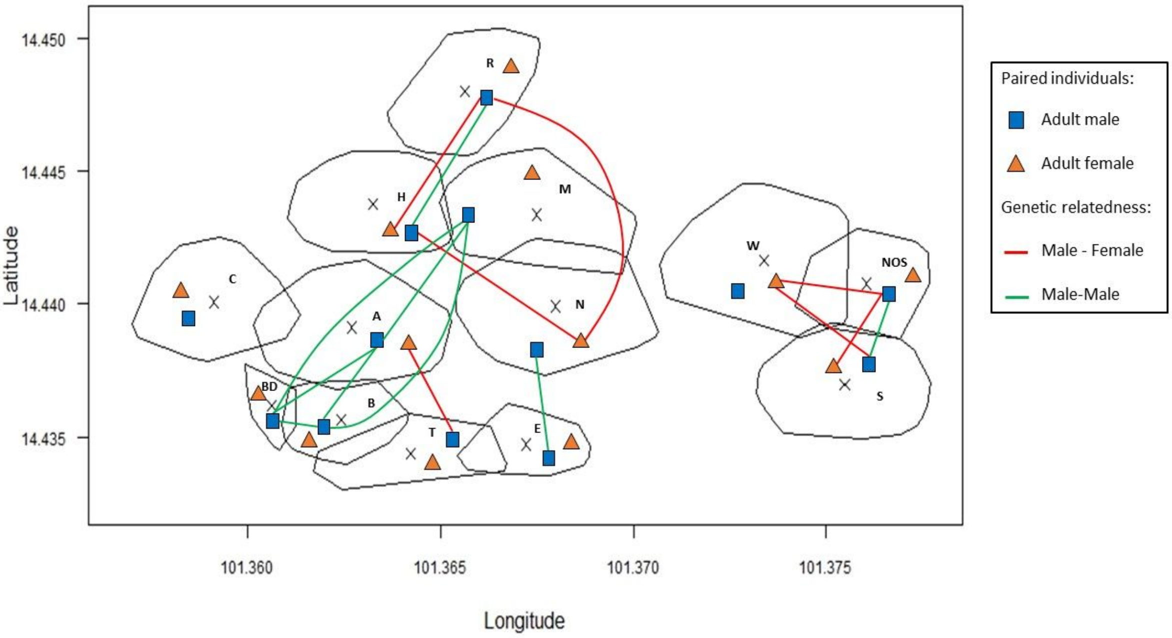

Figure S2.1. Home ranges of 13 habituated white-handed gibbon groups of Mo Singto - Klong E-Tau area, Khao Yai National Park, Thailand (Groups: A, B, BD, C, E, H, M, N, NOS, R, S, T, W) and genetic relationships amongst paired adult males and females. No major changes were observed in territory use and contours during the duration of the study. The map was drawn using 95% kernel density estimation (KDE). X=home range centroid. Territory overlap areas range from 0.12 to 4.66 ha. The genetic relationships between individuals were updated and redrawn from Matsudaira et al. (Matsudaira et al., 2018). Red lines represent the presence of genetic relationships between paired females and paired males and green lines represent the presence of genetic relationships between paired males paired amongst the different study groups. * Kinship: parent-offspring / full-siblings / half siblings. Territory configurations were similar to what has been reported by Asensio et al. (2014) and Savini et al. (2009). Genetic analysis was conducted in 2011 (Matsudaira et al., 2018), with subsequent migration events and changes in group composition documented: In August 2011, female of group W presumably died and was replaced by an adult female dispersing from group S, the daughter of the mated pair. The secondary male in group B emigrated in 2011 to form a new group BD in mid-2012. The secondary male of group A migrated into group B in December 2012. In April 2014, the young adult male of group C migrated into group A as a secondary male and the primary male of group W became secondary male in favour of his son.

Table S2.1 Group compositions and territories area of the white-handed gibbon population at the Mo Singto-Klong E-Tau study site (August 2014)

Group	N individuals	Group composition	Territory area (ha)
A*	3	2AM, 1AF	21.7
B*	5	2AM, 1AF, 1JF, 1I?	14.7
BD	3	1AM, 1AF, 1I?	2.7
C	3	1AM, 1 AF, 1I?	14.77
E	3	1AM, 1AF, 1JM?	12.7
H	4	1AM, 1AF, 1JF, 1I?	16.3
M	5	1AM, 1AF, 1SAF, 1JM, 1I?	18.6
N*	6	2AM, 1AF, 1SAM, 1JF, 1I?	27.5
NOS*	5	2AM, 1AF, 1J?, 1I?	14.1
R	4	1AM, 1AF, 1AF, 1I?	14.2
S	3	1AM, 1AF, 1JM	19.0
T	5	1AM, 1AF, 1SAM, 1JM, 1I?	12.0
W*	4	2AM, 1AF, 1I?	25.0

*Multi-male group; M=male, F=female, ?=sex unknown, A=Adult (age >8years), SA= Sub-adult (5-8 years), J=juvenile (2-5 years), I=infant (<2years).

Table S2.2 Statistical models tested in this study

Model	*Type*	*Response term*	*Test predictors*	*Random terms*	*Sample size*[3]
1	*GLMM*	*Probability of counter-singing*	*Distance between groups*[1]*; I&R group composition*[1]*; MM*[1]*; FM*[1,2]	*I&R*	*N=189; n=13*
2	*LMM*	*Coordination of duets*	*Distance between groups*[1,2]*; I&R group composition*[1,2]*; MM*[1,2]*; FM*[1,2]	*I&R*	*N=189; n=13*
3	*GLMM*	*Probability of pausing for great calls*	*Scenario type; I&R group composition*[1,2]*; Great call position*[1]*; Neighbourhood status*[1,2]	*I&R*	*N=190; n=17*
4	*LMM*	*Initiator group song duration*	*I duet category*[1,2]*; I&R group composition*[1,2]	*I*	*N=389; n=13*
5	*LMM*	*Responder group song duration*	*R duet category*[1]*; Probability of responder group to interrupt*[1]*; I&R group composition*[1]	*I&R*	*N=130; n=13*

[1] *Test predictors fitted in interaction in the initial model;* [2] *Test predictors fitted without interaction in the final model (interaction terms not significant);* [3] *(N=observations; n=groups); all models contained weather as a control predictor; I: initiator duet group, R: responder duet group; analysis were restricted to the first two duet songs (initiator and responder) given during each duet wave, where the second duet was itself not overlapped by a third duet*

Table S2.3 Territorial configuration of the white-handed gibbon population at the Mo Singto-Klong E-Tau study site (August 2014)

Neighbourhood status	Distance between territory centroids (m)			Minimal distance between territory borders (m)		
	min	max	median	min	max	median
Neighbour	406	1201	877	0	290*	0
Non-neighbour	867	3632	2151	177	2865	1224

* For only one neighbouring group pair (H-C), the minimum distance between both territories was not equal to zero due to a lower sampling effort with group C, though we considered this pair as neighbours (see Asensio et al., 2014).

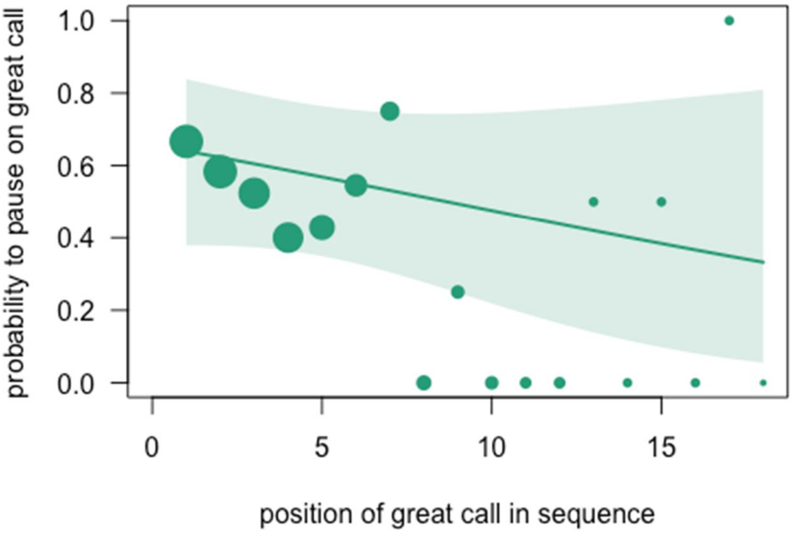

Figure S2.2. Effect of the two-way interaction for scenario 1 only (focal group singing first) and great call position order of focal group produced throughout duets, on the probability of groups to pause for great call sequences of the focal group.

3- WHITE-HANDED GIBBONS DISCRIMINATE CONTEXT-SPECIFIC SONG COMMPOSITIONS

Manuscript published in 2020 in PeerJ 8: e9477 (DOI 10.7717/peerj.9477)

Julie Andrieu[1], Samuel. G. Penny[1,2], Helene Bouchet[3], Suchinda Malaivijitnond[4,5], Ulrich. H. Reichard[6] & Klaus Zuberbühler[1,3]

[1] Department of Comparative Cognition, Institute of Biology, University of Neuchâtel, Switzerland;
[2] School of Pharmacy and Biomolecular Sciences, University of Brighton, UK;
[3] School of Psychology and Neuroscience, University of St Andrews, UK;
[4] National Primate Research Center of Thailand, Chulalongkorn University, Saraburi, Thailand;
[5] Department of Biology, Faculty of Science, Chulalongkorn University, Bangkok, Thailand;
[6] Department of Anthropology and Centre for Ecology, Southern Illinois University, Carbondale, Carbondale, IL, USA;

* Corresponding author: E-mail: j.andrieu.julie@gmail.com (JA)

3.1 Abstract

White-handed gibbons produce loud and acoustically complex songs when interacting with their neighbours or when encountering predators. In both contexts, songs are assembled from a small number of units although their composition differs in context-specific ways. Here, we investigated whether wild gibbons could infer the 'meaning' when hearing exemplars recorded in both contexts (i.e. 'duet songs' vs. 'predator songs'). We carried out a playback experiment by which we simulated the presence of a neighbouring group producing either its duet or a predator song in order to compare subjects' vocal and locomotor responses. When hearing a recording of a duet song, subjects reliably responded with their own duet song, which sometimes elicited further duet songs in adjacent groups. When hearing a recording of a predator song, however, subjects typically remained silent, apart from one of six groups which replied with its own predator song. Moreover, in two of six trials, playbacks of predator songs elicited predator song replies in non-adjacent groups. Finally, all groups showed strong anti-predator behaviour to predator songs but never to duet songs. We concluded that white-

handed gibbons discriminated between the two song types and were able to infer meaning from them. We discuss the implications of these findings in light of the current debate on the evolutionary origins of syntax.

3.2 Introduction

Primate vocal communication is characterised by species-specific repertoires of acoustically distinct vocalisations, some of which are given in response to specific events. The classic example is the vervet monkey (*Chlorocebus pygerythrus*) alarm call system, with acoustically distinct call types given to different predator types (Seyfarth et al., 1980a, 1980b). However, beyond the fact that primate calls can convey relatively distinct meanings, additional complexities have recently come to light, with corresponding implications for evolutionary theories of communication.

First, it is often difficult to characterise a particular call type as an acoustically discrete structural entity. Instead, following in-depth investigation seemingly 'discrete' calls often display considerable amounts of acoustic variation, which may be meaningful to recipients (Keenan et al., 2013). For example, the acoustic structure of chimpanzee (*Pan troglodytes*) rough grunts varies depending on the perceived quality of the food resource (Slocombe and Zuberbühler, 2005), whereas Barbary macaque (*Macaca sylvanus*) barks differ in call duration and mean frequency range according to specific external disturbances (Fischer et al., 1995, 1998; Fischer and Hammerschmidt, 2006).

Second, context can play an important role in how animals interpret each other's calls. Evidence is in terms of how ongoing context modifies how animals react to a specific call type (Zuberbühler, 2000a, 2000b; Arnold and Zuberbühler, 2013; Seyfarth and Cheney, 2018), a mechanism already described by Smith (1977). Empirically, the way intention and external factors affect how primates infer meaning from signals is relatively poorly explored (Grice, 1969; Carnap, 1988; Scott-Phillips, 2010).

Third, call sequences can serve as powerful semantic vehicles beyond the contribution of individual calls (Zuberbühler, 2019a). For instance, the number of roaring units per sequence in guereza colobus monkey (*Colobus guereza*) alarm roars depends on the nature of the danger (Schel et al., 2009). Another example is Campbell's monkeys (*Cercopithecus* *campbelli*) alarm calling, with variation in call rates (Lemasson et al., 2010), call combinations (Ouattara et al., 2009a) and call permutations (Ouattara et al., 2009b, 2009c) depending on external events. Similar phenomena have been observed in putty-nosed monkeys (*Cercopithecus nictitans martini*) (Arnold and Zuberbühler, 2006a, 2006b). Although these findings show remarkable similarities to some aspects of human syntax in terms of combinatorial and permutational properties, the implications for evolutionary theories of language are far from clear, suggesting that more empirical work is needed (Bolhuis et al., 2018; Townsend et al., 2018; Zuberbühler, 2019b).

A relevant primate example of complex combinatorial structure is gibbon song. In most species, mated pairs produce morning duets that appear to serve territorial and mate defence functions (Haimoff, 1984a; Raemaekers and Raemaekers, 1985a; Geissmann, 2002a; Terleph et al., 2015, 2016; Andrieu et al. submitted a). Social learning seems to play some role in the acquisition of song (Koda et al., 2013) and production is subject to social influence (e.g. changes in mating partners usually result in audible differences in song coordination) (Geissmann, 1999b; Terleph et al., 2017). Like most other primate calls, gibbon song contains information about caller identity (Clink et al., 2017; Oyakawa et al., 2007; Terleph et al., 2015) and the caller's physical condition (Barelli et al., 2013a; Terleph et al., 2016). Gibbon songs are audible over long distances, up to 1 km, much beyond an average gibbon home range (Mitani, 1985b), suggesting that they have evolved to communicate to outgroup individuals (Mitani, 1985b; Raemaekers and Raemaekers, 1985a, 1985b; Terleph et al., 2015, 2016).

Interestingly, in white-handed gibbons (*Hylobates lar*), there is also evidence for context-specific song types: duet songs are produced by the mated pair as part of their daily routine while predator songs are given when facing a predator, such as a clouded leopard or python (Clarke et al., 2006; Andrieu et al. submitted b). Both song types are identical in terms of their note repertoires, although there are consistent differences in the prevalence of certain notes and in how notes are combined into songs (Clarke et al., 2006, 2015). Predator songs are sung for longer than duet songs and produced by most group members. They function to deter predators, recruit group members, and alert relatives in adjacent territories (Zuberbühler et al., 1999; Clarke et al., 2006; Matsudaira et al., 2018) . In contrast, duet songs function in mate and territorial defence (Marshall and Marshall, 1976; Raemaekers and Raemaekers, 1985a). Duet songs may also function as indicators of the strength of the social bond of the mated pair, a kind of relationship marker, evidenced by the fact that newly formed pairs appear to go through a lengthy phase of adjusting their relative vocal contributions towards a well-adjusted duet song (Haimoff, 1984a; Geissmann and Orgeldinger, 2000).

Here, we investigated whether gibbons could discriminate the two functionally and structurally distinct song types (i.e. duet song and predator song), by broadcasting natural singing events of a neighbouring group simulated from a concealed speaker. We predicted that if gibbons discriminated between predator and duet songs then they should respond with the matching song types and with behaviour adequate to the situation. Specifically, in response to predator songs we predicted increased vigilance, increased defaecation rates and any other type of anti-predator behaviour already reported in the literature (Boissy, 1995; Clarke et al., 2012). In response to duet songs, we predicted no changes to antipredator behaviour but duet song responses (Raemaekers and Raemaekers, 1985a, 1985b).

3.3 Material and methods

a. Study site and subjects

This study was conducted in the Mo Singto-Klong E-Tau area of Khao Yai National Park, Thailand (101°22′E, 14°26′N), 130 km North-East of Bangkok. Data were collected from December 2012 to August 2014. Thirteen fully habituated groups of white-handed gibbons were monitored, each comprising a primary male, his mated female with her offspring and (in 5 cases) a secondary male, totalling N=53 individuals at the time of the study. Due to a number of constraints, it was only possible to conduct playback experiments with six of the 13 groups (Table 3.1).

Table 3.1. Composition of the study groups at the Mo Singto-Klong E-Tau research area (August 2014).

Group	N individuals	Group composition	Tested	Song provider for
A*	3	2AM, 1AF	-	group H
B*	5	2AM, 1AF, 1JF, 1I?	yes	-
BD	3	1AM, 1AF, 1I?	-	-
C	3	1AM, 1AF, 1I?	-	-
E	3	1AM, 1AF, 1JM	-	-
H†	4	1AM, 1AF, 1JF, 1I?	yes	-
M	5	1AM, 1AF, 1SAF, 1JM, 1I?	yes	group R
N*	6	2AM, 1AF, 1SAM, 1JF, 1I?	yes	group M
NOS*	5	2AM, 1AF, 1J? 1I?	-	-
R	4	1AM, 1AF, 1AF, 1I?	yes	-
S	3	1AM, 1AF, 1JM	-	group W
T	5	1AM, 1AF, 1SAM, 1JM, 1I?	-	group B
W*	4	2AM, 1AF, 1I?	yes	group N

* Multi-male group; M, male; F, female; ?, sex unknown; A, Adult (age > 8 years); SA, sub-adult (5–8 years); J, juvenile (2–5 years); I, infant (<2 years). yes, tested group; -, group not tested.

† No data on latency and duration of first look to speaker due to technical problems (duet playback: female filmed erroneously; predator playback: male moved out of sight).

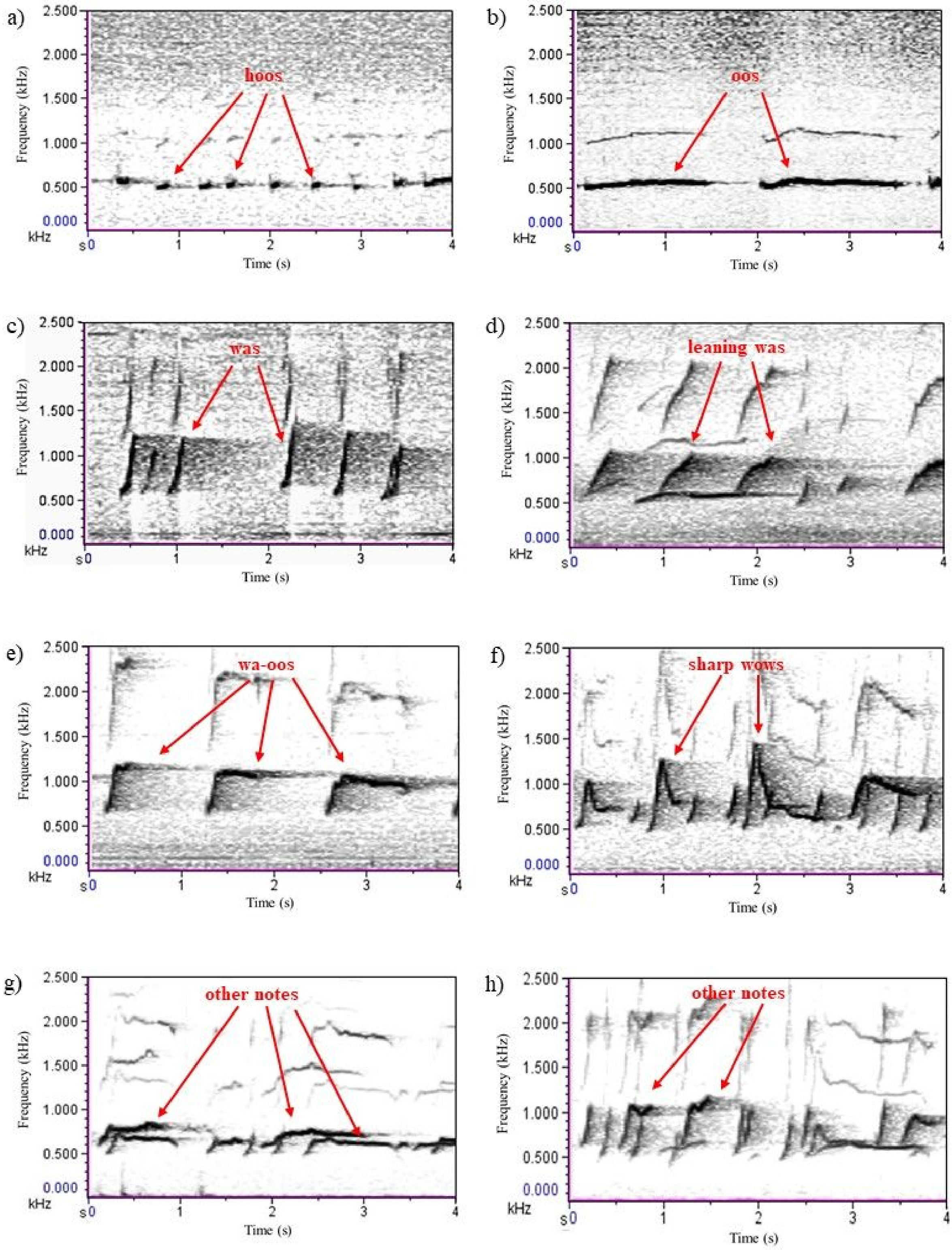

Figure 3.1. Song note repertoire of white-handed gibbons (Raemaekers et al., 1984a, Clarke et al., 2006). Note types a) 'hoo'; b) 'oo'; c) 'wa'; d) 'leaning wa'; e) 'wa-oo'; f) 'sharp wow'; g) and h) 'other notes'. Songs were digitized using Cool Edit Pro 2.1; spectrograms were drawn using 21.6 Hz filter bandwidth, 2.69 Hz frequency resolution, 33.3 ms time grid resolution and a Hanning window function.

b. Terminology

Following Raemaekers, Raemaekers & Haimoff's, (1984a) terminology we distinguished three sequence types within each song: the introductory sequence (series of soft 'hoo' notes, followed by combinations of other note types, such as 'oo', 'wa', 'leaning wa' and 'wa-oo'); the great call sequence (idiosyncratic female call sequence, usually followed by her male's 'coda' response). In duet songs, the first great call sequence usually appears within the first 2 min. Great call sequences can be repeated multiple times (about once every 1–2 min) (Clarke et al., 2006; Raemaekers et al., 1984a; Terleph et al., 2016), in which case they are separated by an interlude sequence (any notes given after a great call sequence, including the final one) (Ellefson, 1968; Raemaekers et al., 1984a) (Figure 3.1 and 3.2a).

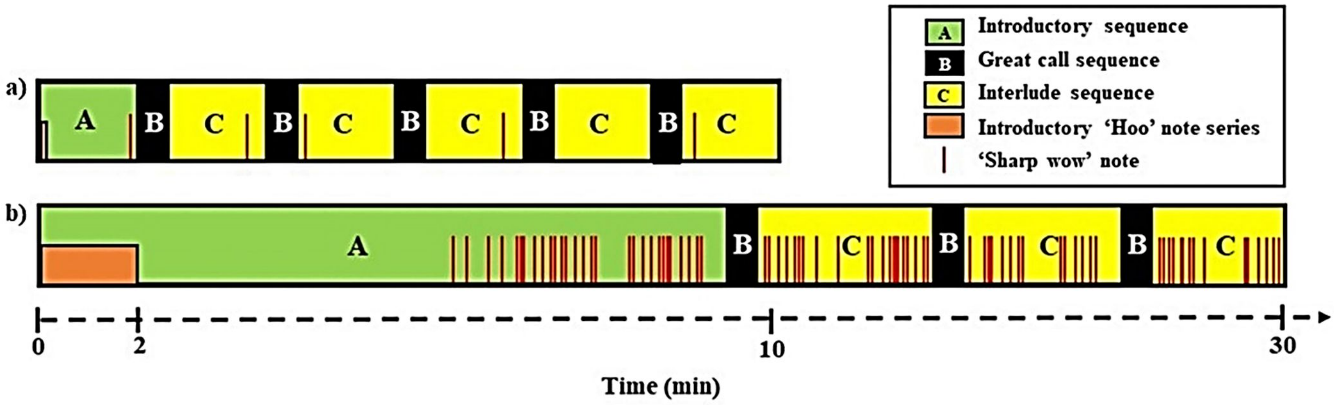

Figure 3.2. Schematic representation of the structural differences between a) duet and b) predator songs (Clarke et al., 2006)**.**

The same three sequence types can also be found in predator songs although, overall, they differ in length and are produced with the contribution of most group members. When comparing predator songs with duet songs, for the introductory sequence the initial 'hoo' notes series are longer and contain more 'hoo' notes, followed by fewer 'leaning wa' notes and more 'hoo' notes (Clarke et al., 2006). The great call sequence is also different, mainly because males respond more rapidly with their answering coda (Clarke et al., 2006). Regarding the interlude sequence, predator songs contain more 'sharp wow' notes, especially towards the end of the song, compared with duet songs(Clarke et al., 2006) (Figure 3.1 and 3.2b).

c. Stimulus collection

Duet songs were recorded on an all-occurrence basis during all-day follows of study groups (Table 3.1) until at least one song suitable as playback stimulus was recorded, that is, a high-quality song with minimum background noise, singing individuals at a maximum distance of 30 m from the recording device. Predator songs were induced by presenting a realistic, life-size clouded leopard (*Neofelis nebulosa*) model to each group following an established protocol (Figure 3.3, Clarke et al., 2006).

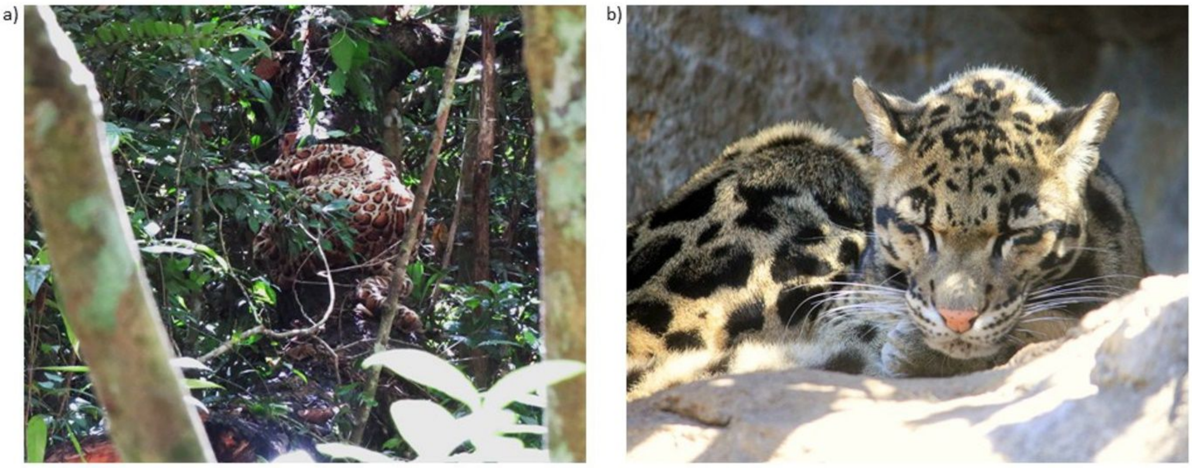

Figure 3.3. a) Clouded leopard model used to elicit predator songs (Photo credit: Julie Andrieu); b) real clouded leopard, *Neofelis nebulosa* (Image credit: goodfreephotos.com at https://www.goodfreephotos.com/animals/mammals/clouded-leopard.jpg.php).

Once a group was located and before positioning the model, we ensured that on the same day the group had (a) already produced at least one duet song more than one hour earlier (to verify a basic motivation to sing), (b) not yet produced a predator song (nor its direct neighbours), (c) not had a natural predator encounter since the beginning of the day-follow, nor heard other species' alarm calls within the last hour and (d) not had an intergroup encounter with a neighbouring group. If these conditions were met, we positioned the predator model on the group's anticipated travel direction outside their visual range. We then continuously recorded their vocal behaviour and scored the presence of any non-vocal anti-predator behaviour on an all-occurrence basis (branch dropping, defaecation, vigilance). Duet and predator songs were recorded using directional microphones (Sennheiser MKH 815T & Sennheiser ME66) with windshields connected to a digital stereo recorder (Marantz PMD660; settings 44.1 kHz, 16 bits) from December 2012 to August 2014.

d. Experimental protocol

Each group was tested once with each stimulus type, which resulted in a total of 12 trials (N=6 duet songs; N=6 predator songs, Table S3.1; minimum interval between trials: 1 week), all broadcasted before 12:00 local time (to match timing of natural duet song production). Prior to playback experiments, we measured the peak intensity of female great call climaxes in spontaneous duet songs (i.e. loudest notes, (Terleph et al., 2016) at an estimated recording distance of 10–20 m using a REED ST-805 (REEDinstruments, Wilmington, NC, USA) sound pressure metre (frequency range 31.5 Hz–8 kHz, measuring level range 30–130 dB, 0.1 dB resolution, accuracy ± 1.5 dB). We measured three great call climaxes per female from the six song-providing groups (Table 3.1), which resulted in a mean sound pressure level of 78.2 ± 8.0 dB (n = 18; dB SPL, A-weighting sound pressure levels for general sound level measurements, and 125 ms fast time weighting). We then broadcasted songs such that subjects always heard recordings from one of their direct neighbours (Table S3.2), with comparable natural audibility (tested at each playback location with a decibel metre, matching climaxes SPL measurements, with real time adjustments in coordination with both experimenters depending on weather conditions on the testing day) and from spatially realistic locations 15–20 m within the canopy from where the song providing group had been seen before within the respective territories.

We standardised the distance between the speaker and subjects to about 150 m (mean ± SD: 149 ± 17 m), with playback conditions randomly counterbalanced (Table S3.2). Stimuli were broadcasted when the same conditions as for predator model presentation had been met, using a Climate CL60-T2 speaker connected to a Kenwood KAC-5203 amplifier, in conjunction with a Roland R-05 digital player.

Playback trials were carried out from spatially realistic locations, that is, from the home range of the song-providing group towards the home range of the target group. In doing so, we took a number of precautions such that the song-providing group could not overhear its own song. Before each trial, we ensured that the song-providing group was not in the vicinity of the speaker (>100 m radius). We then monitored the area for a period of 1 h to further ensure that the song-providing group was not nearby. For each trial, the speaker was positioned in the overlapping zone between the song-providing and target group, such that it was facing away from the home range centre of the song-providing group towards the target group.

Table 3.2. Behavioural response variables extracted for the primary males in both playback conditions

Behaviour	**Definition**
Feeding	Handling or consuming food items
Resting	Prolonged stationary position, with or without eyes closed
Grooming	Auto- or allo-grooming (giver and receiver identity were collected)
Social	Mating, play, aggressive, or parental behaviour
Moving	Travel within or between trees (at least 2 metres)
Vigilance	Scanning the environment, head rotating by at least 45° **(Koenig, 1998)**
Other	Behaviour not classified into any of the above categories
Body Position	**Definition**
Hanging	Suspended in the air, grabbing a branch or a tree part with at least one arm
Sitting/Lying	Sitting on a branch or on the ground / Resting in horizontal position
Gaze direction (staring at a specific location/direction/animal/person for ≥ 3s)	
Speaker	Staring in the direction of the speaker
Ground	Looking towards or actively scanning the ground
Canopy	Looking around, or towards a specific location in the trees at the same elevation as the animal location
Sky	Looking up at the sky
Group member	Looking at a group member (the identity of the receiver was collected)
Observer	Looking at the observer
Elsewhere	Looking in a direction that cannot be classified into any of the above categories
Nowhere	Resting with eyes closed
Other measurements	**Definition**
Elevation (m)	Height of the animal in relation to the ground
Proximity (m)	Distance between the two focal individuals (paired male and female)
Defaecation/Urination	Exuding faeces and/or urine
Dropping branch	Individuals shaking branch(es) so as it ended up falling on
Latency of first look towards the speaker (s)	Time elapsed between stimulus onset and first look towards the speaker
Duration of first look towards the speaker (s)	Duration of first gaze directed towards the speaker location

e. Data Collection

Due to the difficult visual conditions in the forest, it was impossible to continuously videotape the entire duration of trials nor to film all group members simultaneously. We therefore decided to restrict observations to the primary male of each group. Males are easily identifiable by their body hair colouration, facial features and genitals. Primary males were video recorded as long as possible (i.e. until they moved out of sight) using a Panasonic SDR-S26 Camcorder. Videos were coded using ELAN software (ELAN (V5.2) Nijmegen: Max Planck Institute for Psycholinguistics). Because the speaker location was not visible on the video clips (outside camera range) it was necessary for the experimenter to comment on the male's gazing direction during filming, which made blind coding redundant. All video recordings are available on figshare (https://doi.org/10.6084/ m9.figshare.12363050.v1).

Regarding long-term effects, we collected 5-min scan samples of the primary male's behavioural activities, gaze directions, body positions, elevations (m) and proximities to their female partner (m) during 1 h after each trial (i.e. 13 scans per trial; Table 3.2). Furthermore, we scored all defaecation/urination and branch dropping events over a two-hour period using all occurrence sampling

f. Vocal responses

We digitised, analysed and compared songs given in response to both playback conditions, using Raven Pro 64 1.4 (Cornell laboratory of Ornithology, Ithaca, NY, USA). For the introductory sequence, we determined the duration of the initial 'hoo' notes series (s) and the corresponding number of 'hoo' notes, the type of the first ten notes following the 'hoo' series, and the duration of the introductory sequence (i.e. latency to the first female great call). We measured the interval between the female great call and the male coda reply (s), the total song duration (s), and determined whether a neighbouring group also produced a song and its type. Finally, we identified the presence of 'sharp wow' notes and we measured the latency to the first 'sharp wow' note (i.e. time elapsed in seconds between the onset of the song bout and the first 'sharp wow' emitted). This study was approved by the School of Psychology Ethics Committee of St. Andrews University. Approval was given on the understanding that the ASAB guidelines for the Treatment of Animals in Behavioural research and Teaching are adhered to (n°16112011). The research permit was delivered by the National Research Council of Thailand (NRCT, n°0002/5841).

g. Data analysis

Behavioural responses

We compared behavioural responses within subjects and across playback conditions; the primary male's latency and duration of first looks towards the speaker, the occurrence of defecations/urinations and branch droppings, the average distance to their female mate and the canopy heights (medians across all scan samples; Table 3.2). For categorical data (i.e. activity, body position and gaze), we summed up and calculated for each individual the proportion of each behaviour within the categories (see Table 3.2) and compared the behavioural pattern across playback conditions.

Vocal responses

We compared the number of introductory 'hoo' notes and the duration of the introductory 'hoo' notes series, the number of other relevant 'hoo' and 'leaning wa' notes within the first ten notes following the introductory 'hoo' series, and the introductory sequence duration, within groups and across conditions. For the great call sequence, we compared male response delays to the female great calls. Finally, we compared the total song duration between playback conditions, identified the presence of 'sharp wow' notes, and measured the latency to first 'sharp wow' note produced.

Statistical procedures

Due to small sample sizes we opted for non-parametric statistics. Wilcoxon matched-pair signed-rank tests were performed for behavioural data analysis, with exact significance levels reported (Siegel and Castellan, 1988; Mundry and Fischer, 1998). For vocal data, we used Kruskal–Wallis rank sum tests with a Benjamini & Hochberg procedure to correct for multiple testing (Benjamini and Hochberg, 1995). Post-hoc tests were either Wilcoxon rank sum tests with Benjamini & Hochberg p-value adjustments or Dunn (1964)'s tests with Benjamini & Hochberg p-value adjustments for eventual ties. To compare the type of the first 10 notes produced across contexts we used a Pearson's Chi-squared test followed by Chi-squared post-hoc tests with Benjamini & Hochberg p-value adjustments. Statistical analyses were performed using R V3.5.1 (R core Team, 2018) with the significance level set at 0.05.

3.4 Results

a. Vocal behaviour

Response rates

In the duet song condition, 5 of 6 groups responded with duet counter-singing to playbacks of duet songs (Table S3.3). In addition, eight neighbouring groups that shared their borders with the song-providing group or the tested group also produced duet songs during 3 of 6 trials (N=3, N=1, N=4 neighbouring groups, respectively, see Table S3.4), while none of them produced a predator song.

In the predator song condition, 1 of 6 groups responded with a predator song to playbacks of a predator song (within the first 10 min, see Table S3.3). The response song contained a highly delayed first great call and many 'sharp wow' notes, highly typical for a predator song. In addition, two distant (non-neighbouring) groups also produced predator songs during 2 of 6 trials, again characterised by a delayed first great call and 'sharp wow' notes (Table S3.4). None of the groups ever produced a duet song.

Song structure

Playbacks of duet songs reliably triggered synchronised singing by the mated pair of the target groups. To confirm that these vocal responses (N=5) qualified as regular duet songs, we compared them to both spontaneously produced duet songs and experimentally induced predator songs (using a clouded leopard model; Table 3.3) by the same groups. First, there were significant differences across all six variables tested (Table 3.3), while subsequent pairwise comparisons revealed significant differences between predator songs and the two other song types, but not between spontaneous duet songs and response songs elicited by playbacks (Table S3.5 for detailed pairwise comparisons). Second, male latencies to reply to their female's great calls also differed significantly between song types ($\chi^2(2)=33.90$, $P<0.001$, N=82, Kruskal–Wallis rank sum test). Here as well, post-hoc analyses revealed that males gave earlier replies to female great calls in the predatory context (mean delay: −1.7 ± 1.6 s, n=20) than in spontaneous duets (0.6 ± 0.7 s, n=28) or playback duet responses (0.5 ± 0.6 s, n=34) ($P<0.001$ in both cases), with no difference between spontaneous and playback duet responses ($P=0.530$; Dunn's post-hoc test for multiple comparisons, with Benjamini & Hochberg correction).

Table 3.3. Comparison of spontaneous duet songs (*N=5*), predator songs (*N=5*) and songs given in response to playback of duet songs (*N=5*) by the same five groups (Kruskal-Wallis rank sum test).

Variables**	Spontaneous duet song	Predator song	Response song	df	χ^2	P value*
Duration introductory 'hoo' series (s)	8.0 ± 3.1	23.4 ± 6.7	4.7 ± 2.7	2	10.5	<0.05
N introductory 'hoo' notes	11.0 ± 4.5	48.8 ± 14.4	7.4 ± 2.7	2	10.2	<0.05
Song duration (s)	789.4 ± 294.8	2396.4 ± 775.8	1006.8 ± 122.3	2	10.2	<0.05
Latency to 1st great call (s)	101.3 ± 33.5	816.4 ± 368.0	99.0 ± 41.1	2	9.5	<0.05
Latency to 1st 'sharp wow' (s) #	78.1 ± 31.1	370.5 ± 183.2	90.8 ± 35.9	2	9.0	<0.05
N 'sharp wows'	9.2 ± 8.0	362.2 ± 233.9	5.6 ± 6.0	2	9.8	<0.05

Kruskal–Wallis rank sum test for N = 14 songs (W did not produce any 'sharp wow' notes in spontaneous duet; $N_{duet} = 4$, $N_{predator} = 5$, $N_{respons}e = 5$).

* P < 0.05 corrected.

** Means ± SD.

Finally, we compared the first 10 notes produced by males and females immediately following the introductory 'hoo' note series (mean duration: 10.50 ± 2.8 s, n=30, accounting for a total of 100 notes per song type). Significant differences were found between song types regarding their early note composition in 'hoo' and 'leaning wa', but also in 'wa-oo' notes ($\chi^2(4) = 96.86$, P<0.001, Pearson's Chi-squared test). Predator songs contained more 'hoo' notes and fewer 'leaning wa' notes than duet songs, with no differences between spontaneous and playback duet responses. However, 'wa-oo' notes were more common in playback duet responses than spontaneous duet songs, and again in spontaneous duet songs than predator songs (Table S3.6 for detailed pairwise comparisons).

b. Non-vocal behaviour

We were able to record the immediate behavioural responses of primary males in 5 of 6 groups (Table 3.1). All males responded by turning their heads towards the speaker, albeit with no latency differences across playback conditions (median duet: 1.1 ± 1.8 s, predator 2.8 ± 3.3 s, V=2, P=0.188, N_{duet}=5, $N_{Predator}$=5, Wilcoxon matched-pair signed-rank test, Fig. 3.4a). Additionally, we found a trend (although not significant) towards longer gaze duration in the predator than the duet song condition (median duet: 2.0 ± 1.4 s, predator: 12.6 ± 6.1 s, V=0, P=0.063, N_{duet}=5, $N_{Predator}$=5, Wilcoxon matched-pair signed-rank test, Fig. 3.4b).

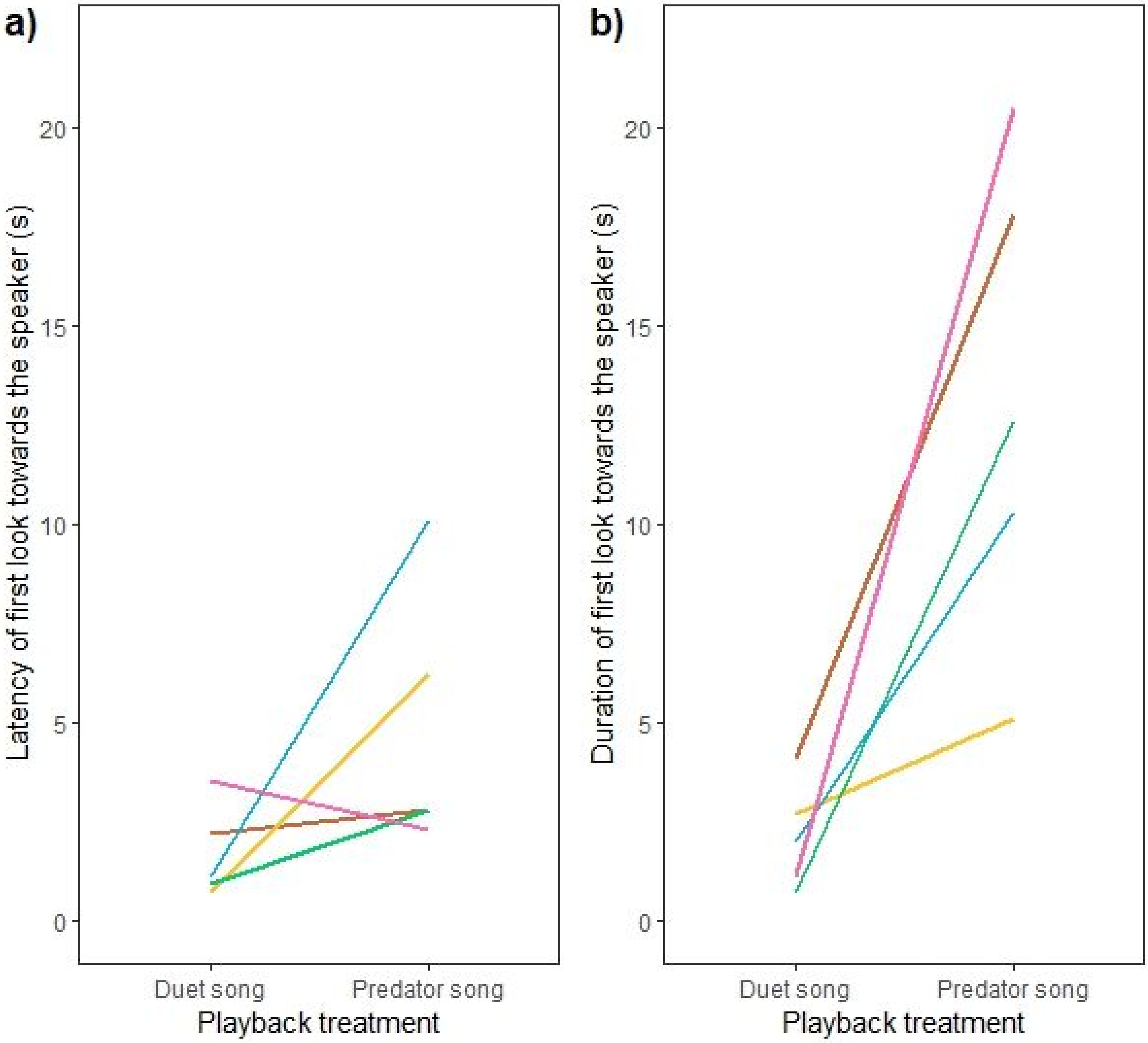

Figure 3.4. a) Latency and b) duration of the male gibbon's first gaze towards the speaker broadcasting a simulated neighbouring group's song (duet vs. predator song condition).

Table 3.4. Comparison of male long-term behavioural responses between playback treatments (Wilcoxon matched pair signed rank tests, *N* = 12 playback trials, with a total of n = 156 scan sampling observations, i.e. 13 scans per individual for 1h).

Variables**		Duet song playback	Predator song playback	V	P value
Behavioural activity	Grooming	1.0 ± 1.6	1.7 ± 2.0	5.5	0.688
	Moving	2.0 ± 1.3	1.7 ± 1.5	13.5	0.594
	Resting	0.8 ± 0.8	0 ± 0	10	0.125
	Feeding	4.0 ± 1.4	0.3 ± 0.8	21	< 0.05*
	Vigilance	2.2 ± 1.3	9.0 ± 2.0	0	< 0.05*
Body position	Hanging	7.2 ± 2.5	5.0 ± 1.7	3.5	0.188
	Sitting/lying	5.8 ± 2.5	8.0 ± 1.7	17.5	0.188
Gaze direction	Speaker	3.0 ± 1.7	4.5 ± 1.1	2	0.125
	Canopy	8.7 ± 1.2	3.2 ± 2.5	21	< 0.05*
	Ground	0 ± 0	5.0 ± 1.3	0	< 0.05*
	Group member	1.3 ± 1.5	0.3 ± 0.8	8.5	0.375
Elevation (m)		17.6 ± 6.2	25.1 ± 7.1	3	0.156
Proximity to mate (m)		8.9 ± 7.3	10.3 ± 7.7	7	0.563
Dropping branch†		0 ± 0	0.5 ± 0.8	0	0.5
Defaecation/Urination†		0.3 ± 0.5	3.2 ± 1.2	0	< 0.05*

† All occurrence behaviours recorded over 2 h post trial

* P < 0.05.

** Means ± SD.

For long-term behavioural responses, we collected data on all six primary males and found no differences across playback conditions in grooming, resting and displacement activities but a significant difference in feeding, with individuals less likely to engage in feeding activities after predator than duet song playbacks (Table 3.4). Regarding anti-predator behaviours, we found no differences in canopy use, distance between mates, and number of branch droppings across conditions.

However, males were more vigilant and defaecated significantly more often following predator compared with duet song playbacks (Table 3.4). Following playback of a predator song, males increased their vigilance activity (Fig. 3.5), directed more gazes towards the ground (Fig. 3.6a) and less towards the upper canopy (Fig. 3.6b) compared with duet treatment (Table 3.4).

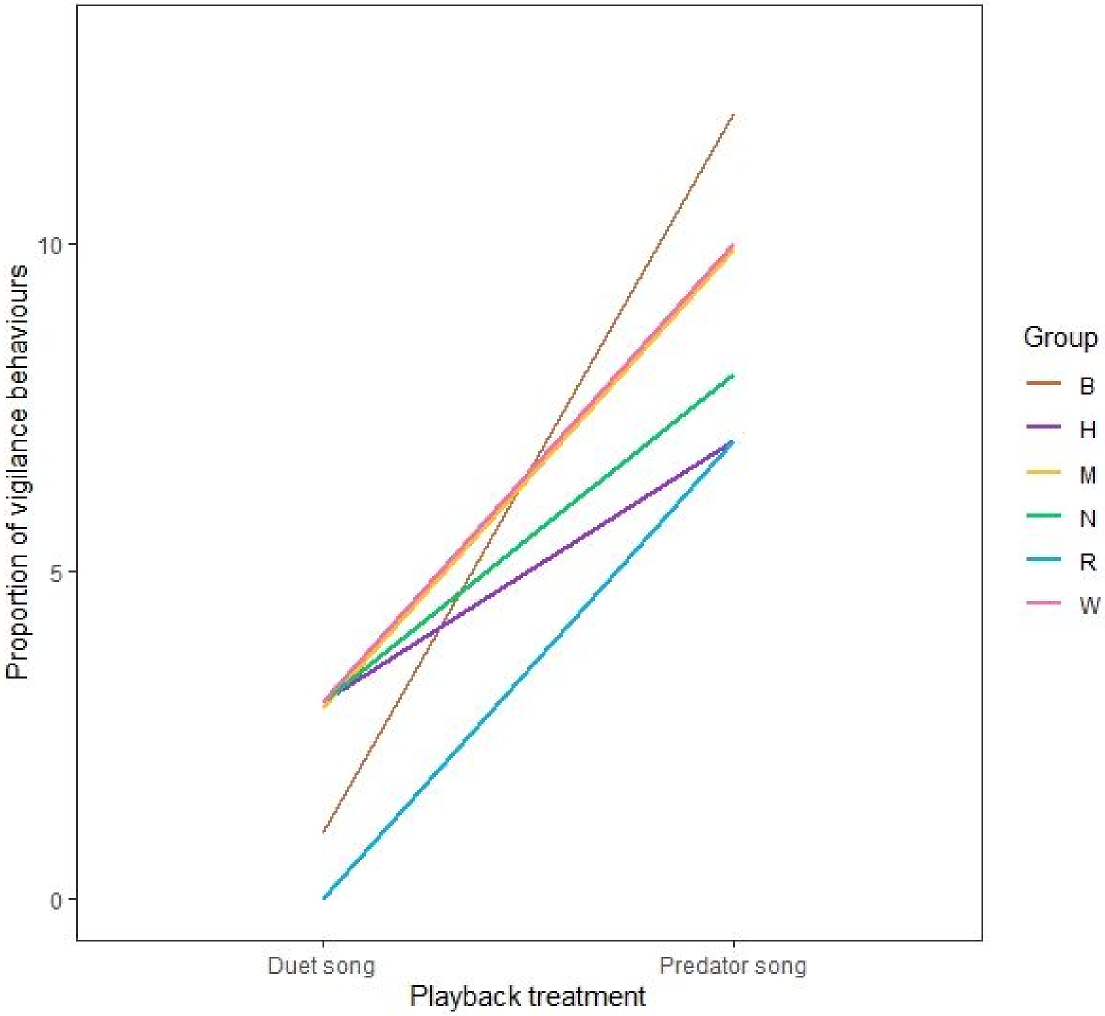

Figure 3.5. Proportion of vigilance behaviours displayed by males in each playback condition (*N*= *6* males).

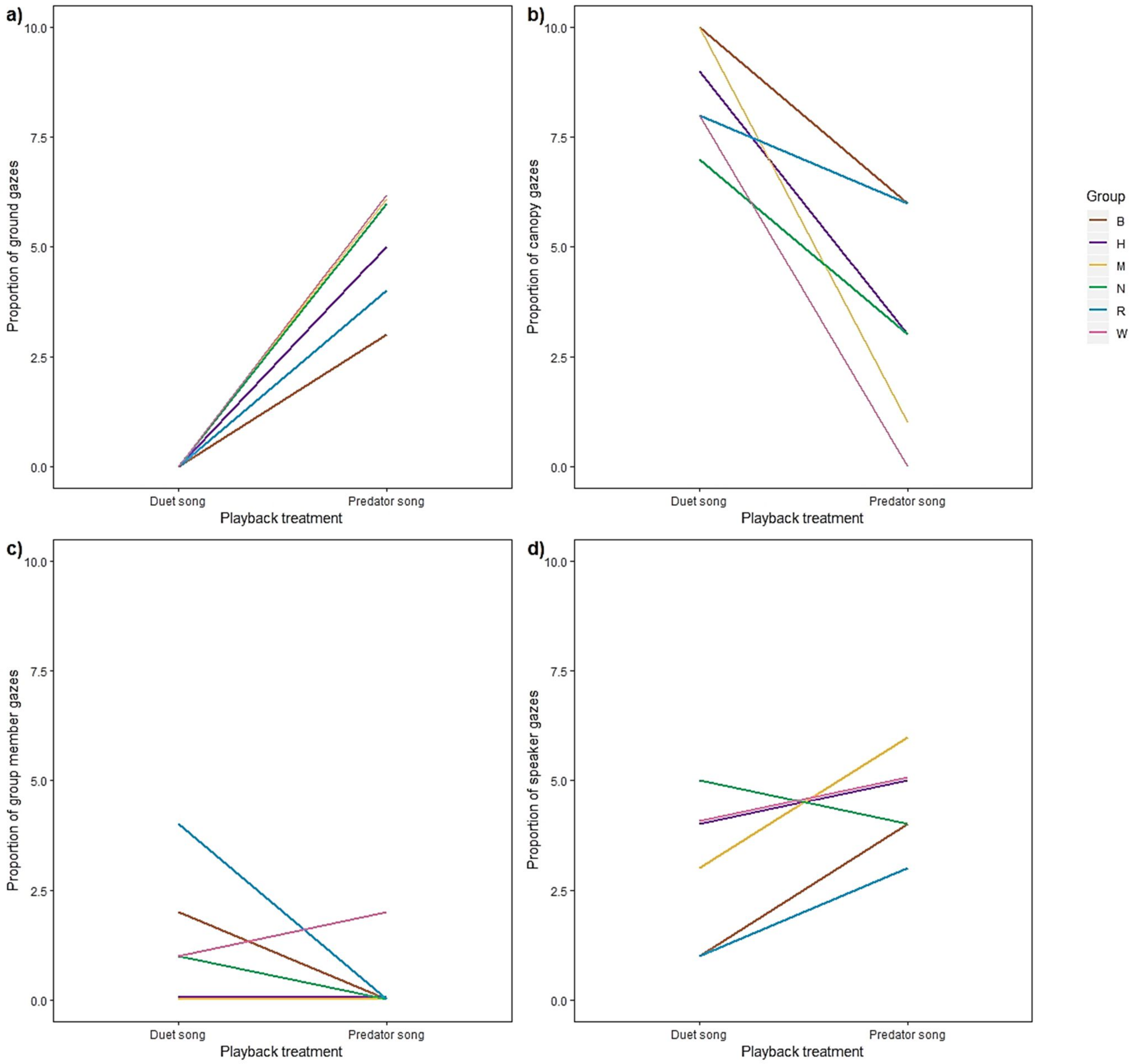

Figure 3.6. Variation of a) ground, B) canopy, c) speaker and, d) group member gazes between playback treatments (*N*= *6* males).

3.5 Discussion

Summary

White-handed gibbons produce two structurally distinct songs in context-specific ways; duet songs (in non-predatory contexts) and predator songs (to clouded leopards and other predators). The two song types differ in the overall duration, frequency and distribution of specific notes ('hoo', 'leaning wa', 'sharp wow') and in the location of the female great calls and male replies within each song. In this study, we investigated whether individuals discriminated between these two structurally different song types and whether they could infer meaning from them. We found several lines of evidence in favour of such an ability. First, playbacks of duet songs reliably elicited natural duet song replies (identifiable by several acoustic parameters) in neighbouring groups and in more distant groups, similar to how natural duet song spread throughout the forest (Raemaekers and Raemaekers, 1985a; Andrieu et al. submitted a). Second, playbacks of predator songs never triggered duet songs in any group, but occasionally predator song replies (identifiable by several acoustic parameters) in one of six neighbouring groups and two non-neighbouring distant groups. Finally, subjects consistently showed anti-predator behaviours (vigilance, ground scanning, defaecation) and a tendency for longer first look towards the speaker after predator compared to duet song playbacks. Based on these data, we concluded that white-handed gibbon song conveys key information about the world, which is made accessible to recipients by a number of structural regularities. This conclusion fits with previous research by Clarke, Reichard & Zuberbühler (2006) who first demonstrated the presence of structural differences in white-handed gibbon songs.

Singing as anti-predator behaviour

Similar to other large cats, clouded leopards are opportunistic predators that attack both terrestrial and arboreal species, including primates (Rabinowitz et al., 1987; Grassman, 2001). Hence, a somewhat surprising finding was that subjects remained mostly silent to others' predator songs, despite showing strong anti-predator behaviour (males and females appeared to behave in the same way, i.e. ground scanning, vigilance, defaecation). The lack of vocal response may be part of a cryptic strategy to conceal the group's location when a dangerous stalking predator is presumed in the vicinity (Aguilar de Soto et al., 2012; Grow, 2019). However, this does not explain why 1 of 6 target groups and two distant groups still responded with predator songs to the playbacks. It is possible that gibbons pursue a flexible vocal strategy, altering between 'crypsis' and 'perception advertisement' depending on perceived personal risk, the ability to benefit neighbouring relatives, and the likely dissuasive effect on the predator itself (Zuberbühler et al., 1999; Clarke et al., 2006).

Equally relevant is the fact that the three predator song responses were shorter than natural predator songs (Tables S3.3 and S3.4). We can think of several explanations for this finding. First, as mentioned already, it is possible that groups tried to minimise their own exposure to the predator if they decided to respond to another group's predator song. Second, differences in predator song duration may function as indicators for perceived urgency, with longer songs indicating more serious threats than shorter songs. We find this less likely to be an evolved function since listeners would have to wait for (and compare) considerable amounts of time periods before extracting the relevant information. Finally, differences in song duration may be linked to how callers perceive the predator (visually, linked to mobbing the predator vs. acoustically, linked to localising the predator). A Direct observation of a real encounter with a tiger is in line with this hypothesis (Uhde and Sommer, 2002). In this instance, group A uncommonly travelled backward towards the tiger's location (spotted 50 m away) and sang for at least 1 h and a half, suggesting that singing primarily serves first and foremost as a predator deterrence device and second as a conspecific warning signal if the exact location of the predator is unknown and groups feel reasonably safe.

Singing as territorial behaviour

In related research (Andrieu et al. submitted a), we have shown that spatial proximity between two neighbouring groups tends to lead to duet song overlap, due to the fact that the second group refuses to delay singing until the first group has finished their duet song. This behaviour is attenuated by kinship, to the effect that related individuals are more likely to respect each other's duets, even if produced at close distances. In the current study, all study groups started producing duet songs while the playback duet song was still being broadcast, suggesting that the manipulation was perceived as a territorial threat. Unfortunately, we could not statistically analyse the effect of genetic relatedness in this study because the sample size was too small (N=6 groups).

Singing as compositional behaviour

Although our study has focussed on song comprehension, it has also generated a more detailed picture of the structural composition of white-handed gibbon songs. Clarke, Reichard & Zuberbühler (2006) already noted that the duet songs of gibbon groups that were not well habituated to human observers contained elements that were normally found in predator songs, notably 'sharp wows'. In our study, all groups were fully habituated to human presence, yet some groups still produced 'sharp wow' notes in their duet song replies to playbacks of neighbouring duet songs, but also in 4 of 5 natural duet songs (Table 3.3), of which 3 were involved in duet counter-signing exchanges with previous duetting direct neighbours. Another structural subtlety concerned the use of 'wa-oo' notes. This note type was near absent in predator songs but common in the early parts of the duet songs, especially the ones given in response to duet song playbacks. We attribute these findings to the fact that our experimental design consisted of playbacks of song recordings at relatively close distances (about 150 m), which may have been perceived as a social threat by some groups, either territorial or risk of partner defection. Future work is required to test whether these notes are actively used to describe events in hierarchically structured ways (main: predatory threat y/n; subsidiary: social threat y/n), similar to how humans represent natural events as tree structures in both cognition and language (Zuberbühler, 2019b).

3.6 Conclusion

Gibbons play an interesting role in questions about the biological roots of language-related capacities in humans. Although part of the Hominoïdae family, they maintain a relatively basal position in their phylogeny by diverging from the great apes some 16 million years ago (Carbone et al., 2014). Nevertheless, gibbons show interesting vocal behaviour by which a small repertoire of acoustically distinct notes are combined into higher-order structures, such as figures, phrases and sequences, assembled into different song types (Raemaekers et al., 1984a; Clarke et al., 2006). These findings have some implications for the ongoing debate about syntax and phonology in animal communication (Bolhuis et al., 2018; Townsend et al., 2018).

In a previous study (Clarke et al., 2006), structural differences between gibbon song types were explained as a case of animal syntax although this was based on a very broad definition of the term. An alternative, more restricted definition of syntax invokes semantics, notably that the units subjected to syntactic operations (e.g. the notes) are meaningful, for which there is currently no evidence in gibbon song. Whatever definition is applied, gibbon song has several levels of complexity and future research should be directed at the acoustic variation in the different note types and their combinations. For example, in the current study we found that the production of ’wa-oo’ and ‘sharp wow’ notes might be linked with perceived social threat. So far, systematic analyses have been restricted to the early parts of the song (based on the assumption that predator information should be conveyed early on) with individual contributions not systematically studied. Traditional acoustic analysis may not suffice to make meaningful progress, suggesting that automated call extraction and categorisation techniques may offer more promise to explore the full combinatorial, hierarchical and compositional capacity of gibbon song (Kershenbaum, 2014; Kershenbaum et al., 2014, 2016; Kershenbaum and Garland, 2015; Fedurek et al., 2016).

3.7 Acknowledgments

We thank the National Research Council of Thailand (NRCT), the National Park, Wildlife and Plant Conservation Department (DNP), and the superintendent of Khao Yai National Park for giving us the permission to conduct this research. We would like to thank Melanie Jackson, Prayoon Saenkhot, Attaklab Chaiyawat, and Surasak Homros for their help in the field. This research project has been funded by the Leverhulme Trust (Research Leadership Award F/00268/AP), the European Research Council (grant number FP7; PRILANG GA283871) and the Swiss National Science Foundation (310030_185324). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

3.8 Supplementary analysis (can be found online at: http://dx.doi.org/10.7717/peerj.9477#supplemental-information.)

Table S3.1. Overview of playback stimuli characteristics ($N_{\text{duet songs}}$=6, $N_{\text{leopard songs}}$=6)

Variables**	Duet stimuli	Predator stimuli
Introductory 'hoo' note series (s)	6.4 ± 8.1	18.8 ± 12.1
Number of introductory 'hoo' notes	6.8 ± 3.7	40.5 ± 823.6
Song duration (s)	1124.9 ± 217.1	2632.7 ± 816.5
Latency to first great call (s)	105.5 ± 20.0	1094.5 ± 650.5
§Latency to first 'sharp wow' (s)	115.4 ± 66.1	354.0 ± 182.2
§Number of 'sharp wow' notes	5.5 ± 7.5	424.0 ± 239.9

§ $N_{\text{duet songs}}$ = 3. Groups A, W and T did not produce any 'sharp wow' note in their duet songs.
** means ± SD.

Table S3.2. Playback experimental design

Trial	Tested group	Stimuli group	Song type
1	W	S	Predator song
2	W	S	Duet
3	R	M	Duet
4	H	A	Predator song
5	B	T	Duet
6	M	N	Duet
7	M	N	Predator song
8	R	M	Predator song
9	H	A	Duet
10	N	W	Predator song
11	B	T	Predator song
12	N	W	Duet

Table S3.3. Overview of group vocal responses to playback treatments

Variables**	**Duet playback** **(N = 5 responding groups)**	**Predator playback** **(N = 1 responding group)**
Latency to reply to stimulus (s)	678.1 ± 70.0	3238.4
Difference between stimulus offset and response onset (s)	-358.5 ± 75.9[#]	580.9
Song duration (s)	1006.8 ± 122.3	1073.5
Latency to first great call (s)	99.0 ± 41.1	819.4
Latency to first 'sharp wow' (s)	90.8 ± 35.9	193.1

[#] Negative values can emerge because all vocal responses to duet playbacks temporally overlapped the stimulus. ** means ± SD.

Table S3.4. Overview of other groups' responses to playback treatments

Variables**	**Duet playback** **(N = 8 responding groups)**	**Predator playback** **(N = 2 responding groups)**
Song duration (s)	768.6 ± 175.8	1425.9 ± 404.5
Latency to first great call (s)	99.0 ± 37.6	493.7 ± 196.9

** means ± SD.

Table S3.5. Comparison of songs given in response to duet playbacks with spontaneous duets and clouded leopard songs given by the same groups (Pairwise comparisons using Wilcoxon rank sum test, with Benjamini & Hochberg corrections).

Context 1	**Context 2**	***P* value**
Introductory ‘hoo’ series duration (s)		
Playback duet response	Spontaneous duet Leopard song	0.151 <0.05*
Spontaneous duet	Leopard song	<0.05*
Number of introductory ‘hoo’ notes		
Playback duet response	Spontaneous duet Leopard song	0.22 <0.05*
Spontaneous duet	Leopard song	<0.05*
Song duration (s)		
Playback duet response	Spontaneous duet Leopard song	0.222 <0.05*
Spontaneous duet	Leopard song	<0.05*
Latency to 1st great call (s)		
Playback duet response	Spontaneous duet Leopard song	0.690 <0.05*
Spontaneous duet	Leopard song	<0.05*
Latency to 1st ‘sharp wow’ (s)		
Playback duet response	Spontaneous duet Leopard song	0.905 <0.05*
Spontaneous duet	Leopard song	<0.05*
Number of ‘sharp wow’ ¤		
Playback duet response	Spontaneous duet Leopard song	0.524 <0.01*
Spontaneous duet	Leopard song	<0.05*

Predator songs were found to be introduced by a longer ‘hoo’ note series, that also contains more ‘hoo’ notes than spontaneous duet songs and duet playback responses (mean ‘hoo’ duration: spontaneous duet: 8.2 ± 3.5 s; playback duet: 4.7 ± 2.7 s; leopard song: 23.4 ± 6.7 s; mean ‘hoo’notes number: spontaneous duet: 11.0 ± 4.5; playback duet: 7.4± 2.7; leopard song: 48.8 ± 11.4). Furthermore, predator songs were found to be longer in duration with a delayed first great call production compared to spontaneous duet songs and duet playback responses (mean song duration: spontaneous duet: 794.5 ± 340.1 s; playback duet: 1006.8 ± 122.3 s; leopard song: 2396.4 ± 775.8 s; mean latency to first great call: spontaneous duet: 104.8 ± 37.6 s; playback duet: 99.0 ± 41.1 s; leopard song: 816.4 ± 368.0 s).

Additionally, differences emerged when analysing the production latency of the first ‘sharp wow’ and in the number of ‘sharp wow’ notes produced, with predator songs containing more ‘sharp wow’ notes with a delayed production (mean latency to first ‘sharp wow’: spontaneous duet: 78.1 ± 31.1 s; playback duet: 90.8 ± 35.9 s; leopard song: 370.5 ± 183.2 s; mean ‘sharp wow’ notes number: spontaneous duet: 11.5 ± 7.1; playback duet: 5.6 ± 6.0; leopard song: 362.2 ± 233.9).

Table S3.6. Comparison of the first ten notes produced across singing contexts (duet playback responses, spontaneous duets and clouded leopard songs given by the same five groups, N=15) (Pairwise comparisons using Chi-squared post-hoc tests, with Benjamini & Hochberg corrections, n=300 notes).

Context 1	**Context 2**	*P* value
'hoo' notes		
Playback duet response	Spontaneous duet Leopard song	1 <0.001*
Spontaneous duet	Leopard song	<0.001*
'leaning wa' notes		
Playback duet response	Spontaneous duet Leopard song	1 <0.05*
Spontaneous duet	Leopard song	<0.05*
'wa-oo' notes		
Playback duet response	Spontaneous duet Leopard song	<0.001* <0.001*
Spontaneous duet	Leopard song	<0.001*

(* P<0.05).

Leopard songs were found to contain more 'hoo' notes and less 'leaning wa' than spontaneous duets and playback duet songs, with no differences between spontaneous duets and playback duet songs ('hoo': spontaneous duet: 1.2 ± 1.3; playback duet: 1.2 ± 1.5; leopard song: 6.1 ± 2.5; 'leaning wa': spontaneous duet: 1.0 ± 1.6; playback duet: 1.0 ± 1.3; leopard song: 0.2 ± 0.6).

However, songs in response to duet playbacks contained more 'wa-oo' notes than spontaneous duet songs and predator songs, with spontaneous duet songs containing also more 'wa-oo' notes than predator songs ('wa-oo': spontaneous duet: 1.8 ± 2.0; playback duet: 5.0 ± 2.3; leopard song: 0.1 ± 0.3).

4- GIBBON SONGS REFER TO EXTERNAL EVENT

Manuscript submitted to Animal Behaviour

Julie Andrieu[1], Samuel. G. Penny[1,2], Helene Bouchet[3], Suchinda Malaivijitnond[4,5], Ulrich. H. Reichard[6] & Klaus Zuberbühler[1,3]

[1] Department of Comparative Cognition, Institute of Biology, University of Neuchâtel, Switzerland;

[2] School of Pharmacy and Biomolecular Sciences, University of Brighton, UK;

[3] School of Psychology and Neuroscience, University of St Andrews, UK;

[4] National Primate Research Center of Thailand, Chulalongkorn University, Saraburi, Thailand;

[7] Department of Biology, Faculty of Science, Chulalongkorn University, Bangkok, Thailand;

[8] Department of Anthropology and Centre for Ecology, Southern Illinois University, Carbondale, Carbondale, IL, USA;

* Corresponding author: E-mail: j.andrieu.julie@gmail.com (JA)

4.1 Abstract

White-handed gibbons produce structurally complex, hierarchically organised songs by repeating basic notes at different rates and combining these basic notes into figures, phrases and sequences, often as part of finely tuned vocal interactions. Singing can be part of competitive interactions with neighbours or part of anti-predator behaviour, notably to clouded leopards and reticulated pythons. Given the different implications of these situations, we were interested in whether gibbon songs contained sufficient structural difference for listeners to draw inferences about the associated external event. In field experiments, we systematically exposed eleven groups to authentic predator models and compared their behavioural and vocal responses between conditions. We found consistent differences at all levels of song organisation, i.e. at the note, figure, phrase and sequence level, especially during the early parts of song. We concluded that white-handed gibbon song evolved as a potent communication device that allows signallers to reliably convey external events, allowing listeners to make inferences about relevant events.

4.2 Introduction

Alarm calling is a common anti-predator strategy that has evolved in large parts of the animal kingdom. In primates, alarm calls are usually given by individuals after detecting a predator, such as felids, raptors or snakes, often in a dual function to both warn conspecifics and also deter the predator (Zuberbühler et al., 1997, 1999; Caro, 2005). From a cognitive viewpoint, animal alarm calls are remarkable because, similar to human language, they often appear to refer to specific events in the environment which then triggers specific behavioural responses in recipients. The classic example is vervet monkey (*Chlorocebus pygerythrus*) alarm calls, with adults producing acoustically distinct vocalisations to aerial and terrestrial predators and recipients responding in appropriate ways (Seyfarth et al., 1980a, 1980b). Similar findings have been reported in many other species, such as mustached tamarins (*Saguinus mystax*) (Kirchhof and Hammerschmidt, 2006), Verreaux' sifakas (*Propithecus verreauxi*) (Fichtel, 2008), Campbell's monkeys (*Cercopithecus campbelli*) (Ouattara et al., 2009a), Diana monkeys (*Cercopithecus diana*) (Zuberbühler, 2000c, 2001) or titi monkeys (*Callicebus nigrifrons*) (Cäsar and Zuberbühler, 2012).

More recently, it has become clear that producing acoustically distinct calls is not the only way by which animals communicate about external events. For example, some alarm call systems are based on acoustically graded features and differences in call intensity (e.g. chacma baboons, *Papio cynocephalus ursinus*, Fischer et al., 2001a, 2001b). Other systems are based on differences in call rates. For example, Guereza colobus monkeys (*Colobus guereza*) specify the nature of the perceived danger (eagle vs. leopard) by altering the number of calls per sequence (Schel et al., 2009) whereas Campbell's monkeys modulate call rate depending on urgency (Lemasson et al., 2010). Finally, there is also evidence for context-specific combinations of calls encoding the nature of the danger in putty-nosed monkeys (*Cercopithecus nictitans*) (Arnold and Zuberbühler, 2006) and Campbell's monkeys (Ouattara et al., 2009b, 2009c) or even revealing the location of the predator in titi monkeys (Cäsar et al., 2013).

Despite this wealth of research, there is peculiar imbalance in the distribution of evidence across the various groups of primates, with some groups (such as macaques or great apes) contributing little or nothing to the overall dataset, despite hundreds of years of combined observations. Gibbons are part of this underrepresented group, with close to no information available on alarm calling behaviour, with the exception of white-handed gibbon (*Hylobates lar*) song, which are rich and structurally complex compositions of different notes assembled into higher order structures, such as ‘figures’, ‘phrases’ and ‘sequences’ (Raemaekers et al., 1984a). Gibbon duet songs have mainly been studied in their function in intergroup competition over mating opportunities and resources (Raemaekers and Raemaekers, 1985a; Terleph et al., 2015, 2016). A recent finding in line with this research is that taking turn in duet interactive singing between groups is enhanced by the presence of genetic relatives (Andrieu et al. submitted a). Although gibbons are reasonably well protected from ground predators, due to their preferential use of the higher canopy strata and highly specialised locomotion, they have a relatively small body size compared to other Hominoïdae, which makes them more vulnerable to predators. In an earlier study, it was found that songs given to predators were composed of almost the same units as regular duet songs but arranged in different ways (Clarke et al., 2006). More specifically, predator songs contained more ‘sharp wow’ notes, were longer, with delayed great call sequences, and produced by more group members (Clarke et al., 2006). Whether gibbon songs qualify as referential alarm call system, as found in vervet monkeys, is currently unknown. While natural encounters with reticulated pythons reliably triggers predator songs (Neudenberger, 1993; Uhde and Sommer, 2002) (Andrieu personal observation), experiments with a python model have failed to trigger the consistent anti-predator responses (Clarke et al., 2006, 2012), suggesting that the previous python model was not sufficiently representative of real encounters, and that gibbons might have very sensitive perception capacities concerning predator recognition.

Here, we revisited the issue by testing whether the predator songs of white-handed gibbons were structurally distinct when encountering pythons or leopards. To this end, we fabricated authentic clouded leopard model and new reticulated python model as close as possible to a real animal (see supplementary Figure S4.1), which we presented in systematic ways to eleven habituated groups of wild white-handed gibbons.

4.3 Material and methods

a. Study site and subjects

This study was conducted in the Mo Singto-Klong E-Tau area of Khao Yai National Park, Thailand (101°22'E, 14°26'N), 130 km North-East of Bangkok. Khao Yai National Park was established in 1962 and covers a total area of 2,165 km^2. The Mo-Singto – Klong E-Tau research site is located in the central sector of the park at an elevation of 730-870 m above sea level, with an average annual rainfall of 3,000 mm, mostly during a wet season from May through October (Whitington and Treesucon, 1991).

Data were collected from December 2011 to August 2014. Thirteen habituated groups of white-handed gibbons were monitored. Habituation started in the 1980s, with long-term demography data available for all groups, including the genetic relationships between most individuals (Reichard, 2009; Matsudaira et al., 2018). Predator experiments were conducted with 11 out of the 13 groups totalling N=45 individuals (the two remaining groups were not fully habituated to human presence, see supplementary Table S4.1).

b. Terminology

Gibbon song has a complex, hierarchically organised, species-specific structure. The basic acoustic units are called *'notes'*, which, in white-handed gibbons, have a frequency distribution between 400-1,600 Hz with ascending frequency transitions. At least eight types of notes can be identified (see supplementary Figure S4.2; Raemaekers et al., 1984a: 'oo', 'wa', 'leaning wa', 'wa-oo', 'ooaa', 'sharp wow'; Clarke et al., 2006: 'hoo' and other notes).

'*Figures*' are a distinctive assemblage of single notes uttered in series, or a combination of a single note and an exhaling event, three of which have been described in white-handed gibbons: *trills* (males only), *Ooaa bursts* and *Hoo-oos*. A male *trill* is composed of 'wa' notes series produced in one exhalation, *Ooaa bursts* are composed of a series of 'Ooaa' notes and *Hoo-oo figures* are composed of both inhalation and exhalation events; the exhaling note can be either a short 'hoo' or a long 'oo' note (Figure S3.3.a, Raemaekers et al., 1984a).

'*Phrases*' comprise a 'larger and looser' succession of notes separated by short pauses and are produced in an idiosyncratic manner by a single individual. The most recognisable and stereotypes 'phrase' is the female great call, which starts with a series of female 'hoo' notes, during which the male usually falls silent, followed by the female proper great call phrase. A great call lasts from 12 to 22s and contains 6-13 notes with increasing frequency from 400-1,200Hz, followed by a decrease to 400Hz, a second increase to 1,600 Hz, the climax phase, and a subsequent drop to 400 Hz (Figure S4.3.b).

‘*Duet songs*’ are given by the mated pair in a synchronised fashion in non-predatory contexts with male notes being typically higher pitched than female notes (Marshall and Marshall, 1976; Haimoff, 1983; Raemaekers et al., 1984a). Duet songs are composed of three sequences types: the introductory sequence, starting with a series of soft ‘hoo’ notes followed by other notes types (i.e. ‘oo’, ‘wa’, leaning wa’ and ‘wa-oo’ notes); the great call sequence, which comprises the female great call phrase and her mate’s answering coda (Figure S3.3.b), and can be repeated multiple times about once every 1-2 min (Raemaekers et al., 1984a; Clarke et al., 2006) and the interlude sequence, defined as all notes given between great sequences and after the last great call sequence (Ellefson, 1968; Marshall and Marshall, 1976; Haimoff, 1981; Raemaekers et al., 1984a).

Predator songs are typically given to predators such as tigers, clouded leopards, reticulated pythons, and humans (unhabituated groups only) (Clarke et al., 2006). They are composed of the same three sequence types as duet songs, the same songs units, but they differ in some structural ways. Overall, predator songs are longer and contain more ‘sharp wow’ notes that are produced in the later part of the song (Clarke et al., 2006). The introductory sequence opens with a longer soft ‘hoo’ notes series and ends with a highly delayed great call sequence production in which males were found to respond more rapidly to their mate (Clarke et al., 2006). Also, a number of incomplete song types have been described, e.g. ‘soft hoo-oo songs’ (composed of soft ‘hoo’ and ‘oo’ notes only) to pythons, exclusively, but these utterances are not comparable to the full predator and duet songs (Clarke et al., 2006).

c. Experimental procedure

Natural predator encounters are challenging to witness in dense forest habitats. Over 28 months of field research we only witnessed two natural encounters reticulated with reticulated pythons (*Python reticulatus*) but none with clouded leopards (*Neofelis nebulosa*) (see supplementary section). A common solution to study anti-predator behaviour systematically is to present realistic predator models in controlled ways. In this study, we presented models of a clouded leopard and a reticulated python to different groups, following a previously established protocol (Clarke et al., 2006). For the leopard, we used a model already used by

Clarke et al. (2006, 2012), made of commercially available leopard print fabric wrapped around a large backpack, which previously elicited strong anti-predator behaviours such as dropping from canopy height, increasing vigilance, and multiple defecation as a sign of distress (Boissy, 1995; Clarke et al., 2006, 2012). For the python, we fabricated a life-size, realistic model by copying the different body parts from pictures of reticulated pythons onto textile fabric with thermo-active transfer paper (www.avery.fr/transfert-textile, Avery ©). The different parts were then sewn together and filled with synthetic fibre to produce a realistic reticulated python model.

d. Data collection

We located groups after they had produced their first morning duet song. Before positioning a predator model, we ensured that: a) there was a delay of at least 1 hour since the last duet song, b) no predator song had been produced by the target group and its neighbours prior to model presentation, c) no natural predation event, intergroup encounter or alarm calls given by other species had happened before model presentation.

When these conditions were fulfilled, we positioned the predator model and removed ourselves from the groups' line of sight. Eleven groups were tested on different occasions with both models between 2011-2014 (2011-12: N=9 leopard by HB and JA; 2013-14: N=11 python and N=2 leopard by JA and SP). Duet songs were opportunistically recorded between 2013-2014 by JA and SP. Group compositions for all 11 groups were identical for all three conditions.

To compare songs produced in the three conditions, we continuously recorded the entire singing events, using Sennheiser MKH 815T and ME66 directional microphones with windshields connected to a Marantz PMD660 digital recorder set at a 44.1 kHz/16 bits sampling rate. Whenever a song was produced, we determined a) caller identity; b) song type (duet or predator); c) song duration (s); d) total number of great calls and production latency relative to song onset; and e) any relevant external event. During predator model presentations, we additionally collected the identity of the individual(s) that spotted the model first, as well as the detection time preceding the onset of the song.

e. Data analysis

Non-vocal responses

For the python model, we recorded the behaviour of all focal animals (N= 11 primary males, N=11 females; Table S4.1), from one hour before (range 30-60 min) until one hour after detecting the model (range 30-60 min). We used scan sampling at five-minute intervals for behavioural activity, gaze direction and body position (Table S4.2). We used all occurrence sampling for defecation/urination behaviour and branch dropping, again one hour prior and after model detection. We compared, for each individual, the median distance to its mate and elevation in the canopy, as well as the median defecation/urination and branch dropping rates (Table S4.2), for the same time period before and after model detection (60-120min total). For categorical data, we compared, for each individual, the proportion of each behaviour per category (Table S4.2), before and after model detection. For gaze, we compared, for each individual, the proportion of gazes directed to the ground and towards the predator model, again before and after detection. We used non-parametric Wilcoxon signed-rank tests for null-hypothesis testing. Since group members are likely to influence each other in their responses, we used a permutation test by running the exact Wilcoxon signed-rank tests with 1,000 iterations, selecting randomly one individual per mated pair among the 11 groups tested. Note that comparable analyses for the clouded leopard model have already been conducted and published (Clarke et al 2006, 2012).

Vocal responses

We randomly selected one duet song per group from a large sample collected during daily follows (N=11). Predator models (i.e. leopard and python) were only presented once to each group resulting in N=22 predator presentation recordings. We digitised, analysed and compared songs using Raven Pro 64 1.4 [Cornell laboratory of Ornithology, Ithaca, NY]. As we could not perform a discriminant function analysis (DFA) given our limited sample size (as the minimum number of songs required would be 4 to 5 times the number of predictor variables), we compared variables across contexts (i.e. duet, python and leopard songs) using either linear mixed models (LMMs) or conventional non-parametric tests.

At the *note level*, we extracted, for each song, the first five clearly audible 'hoo' notes produced by the mated pair. For predator songs, we extracted 'hoo' notes only from individuals who had spotted the model. We measured the duration (s), low, peak and delta frequency (Hz) for all 'hoo' notes, as well as their inter-call intervals. For the first five 'hoo' notes, we used five

LMMs (one per acoustic parameter) to perform pairwise comparisons with context (three-level) and sex (two-level) as predictor variables and individual identity as random intercept. We included a two-way interaction between context and sex (as males are known to be higher pitched than females; Clarke et al., 2015). Prior to fitting, LMM response variables were Yeo-Johnson transformed (Weisberg, 2001) and Shapiro-Wilk tested for normality, and residuals from the full models were visually inspected for normality and homoscedasticity. We compared full and null models with a likelihood ratio test (LRT) and removed non-significant interaction terms (Engqvist, 2005). We performed pairwise comparisons between contexts based on difference of least-squares means (Lenth, 2016) using Bonferroni p-value corrections. We also measured the latency to first 'sharp wow' note from song onset (s) and carried out comparisons across contexts using a Kruskal-Wallis rank sum test followed by a Dunn (1964)'s post-hoc test for multiple comparisons with the Bonferroni correction. We further categorised each song as a function of 'sharp wow' notes being common (≥25) or rare (<25) and used Fisher's Exact tests for count data (two-sided) and the associated Bonferroni-corrected pairwise comparisons across contexts.

At the *figure level*, we counted the number of *Hoo-oo figures* present in the initial song segments exclusively composed of 'hoo' and 'oo' notes (until the first non 'hoo' and 'oo' note is produced) and measured the duration of the inhaling event (s) (see Figure S4.3.a). As duets do not contain *Hoo-oo figures*, we compared the number of *Hoo-oo figures* in python and leopard songs, using exact test of differences in Poisson rates. We then compared the duration of inhaling events across contexts with a Mann-Whitney test (two-tailed).

At a *higher structural level*, we focussed on the early song segments that open the introductory sequence, predicting that any information about external events should be encoded as early as possible. We measured the duration of the opening 'soft hoo' notes series (s) (i.e. from song onset to the first non 'hoo' note produced); the following 'oo' segment composed of 'soft oo' and 'hoo' notes (s) (i.e. from the first 'oo' note produced until the production of the first non 'hoo' and 'oo' note); and the combination of these two segments(the 'hoo-oo' notes segment, from song onset to the first non 'hoo' and 'oo' note) . Song segments were compared across contexts again using LMMs with context as predictor variable (three-level) and group identity as random factor, followed by subsequent pairwise comparisons (least-squares means (Lenth, 2016) using Bonferroni p-value corrections). Prior to fitting, LMM response variables were Log-transformed and Shapiro-Wilk tested for normality, and residuals were checked for normality and homoscedasticity.

At *sequential levels*, we measured the duration between the female great call and her mate's coda and non-coda replies (s) (Figure S4.3.b) and explored male reply latency in the different singing contexts using a Kruskal-Wallis rank sum tests followed by Bonferroni-corrected pairwise comparisons. We also compared male response type (coda vs. song follow-up) across conditions using a Fisher's Exact test for count data (two-sided) followed by the associated pairwise comparisons.

Lastly, we determined the duration of the introductory sequence (from song onset to the first great call sequence) and the total song duration (s) across contexts, again by Log-transforming variables to achieve normality, followed by LMMs and subsequent pairwise comparisons.

All statistical analyses were performed using **R** v3.5.1 (R core Team, 2018).

4.4 Results

a. Overall response rates

We witnessed two natural encounters with a reticulated python, but no encounter with a clouded leopard during the entire study period. In both cases, the groups showed strong anti-predator behaviours and sang ('soft Hoo-oo song' and full predator song; see supplementary section). In response to the python model, 10 of 11 groups sang in conjunction with various anti-predator behaviours compared to the 3 out 9 groups that sang back to the old python model (Clarke et al., 2006). 1 of 11 groups detected but ignored the model and carried on with its daily routine. 11 of 11 groups showed strong anti-predator behaviours and sang to the clouded leopard model.

b. Non vocal behaviour

Upon detecting the python model, gibbons significantly dropped in canopy height, increased their vigilance level, scanned more actively the predator model and increased their defecation/urination, with no differences in distance between the mated pair and branch dropping rate (Table 4.1, for full behavioural results see Table S4.3). Results are in line with what was previously found for leopard model presentations (Clarke et al., 2012, see methods).

Table 4.1 Anti-predator behaviour of white-handed gibbons to python model

Variable	Detection		Statistics	
	Before Mean ± SD	After Mean ± SD	Corrected V value†	Corrected P value†
Elevation (m)	25.2 ± 6.0	14.5 ± 6.4	54.6	< 0.01*
Proximity (m)	12.9 ± 11.8	9.6 ± 9.3	26.1	0.718
Vigilance	0.1± 0.3	6.1 ± 3.2	0	< 0.01*
Gaze towards ground/predator	0.1 ± 0.3	6.0 ± 3.3	0	< 0.01*
Defecate/urinate	0.2 ± 0.4	1.4 ± 0.9	0	< 0.05*
Branch drop	0 ± 0	0.6 ± 0.9	0	0.135

N=11 model presentations, with a total of 376 scan sampling observations for 22 individuals, i.e., 188 before-after comparisons per condition. Exact Wilcoxon signed-rank Permutation test with 1,000 iterations. Mean behaviours occurrence was calculated over all scan samples; †Sample size reduced to 188 scan sampling observations for 11 individuals when correction was applied. (* P < 0.05).

c. Vocal behaviour

Response rates

All groups produced long and loud predator songs to the clouded leopard model (N= 11). The female of one group (T) did not produce great calls during the leopard song, and this song could not either be qualified as a 'Hoo-oo song' (see methods), so we excluded this group from the analysis targeting great calls only. Seven of ten groups produced long and loud predator songs to the reticulated python model. Three of ten groups produced less conspicuous songs (i.e. 'Hoo-oo songs'), composed only of soft 'hoo' and 'oo' notes, but no great calls, which we also excluded from the subsequent analyses. For the N=18 songs that were retained for acoustic analyses, we checked whether they fitted previous descriptions of predator song following Clarke et al. (2006) (song duration, latency to first great call, number and latency to first 'sharp wow').

Acoustic structure of notes

We first compared the acoustic structure of male and female 'hoo' notes across contexts (duet: N=78, python: N=57, leopard: N=97). For duration, the full model differed significantly from its null model ($\chi^2(5)=24.83$, $P<0.001$) with a significant two-way interaction between context and sex before (LRT, $\chi^2(2)=15.92$, $P<0.001$) but not after correction (Figure 4.1a, Table S4.4 for detailed pairwise comparisons). For peak frequency, the full model differed significantly from its null model ($\chi^2(5)=96.43$, $P<0.001$) with a significant two-way interaction between context and sex (LRT, $\chi^2(2)=7.87$, $P<0.05$), due to males producing higher pitched 'hoo' notes than females in similar singing context (Table S4.4), with both sexes producing higher-pitched 'hoo' notes in the duet than predatory contexts with no difference between leopard and python songs (Figure 4.1b, Table S4.4).

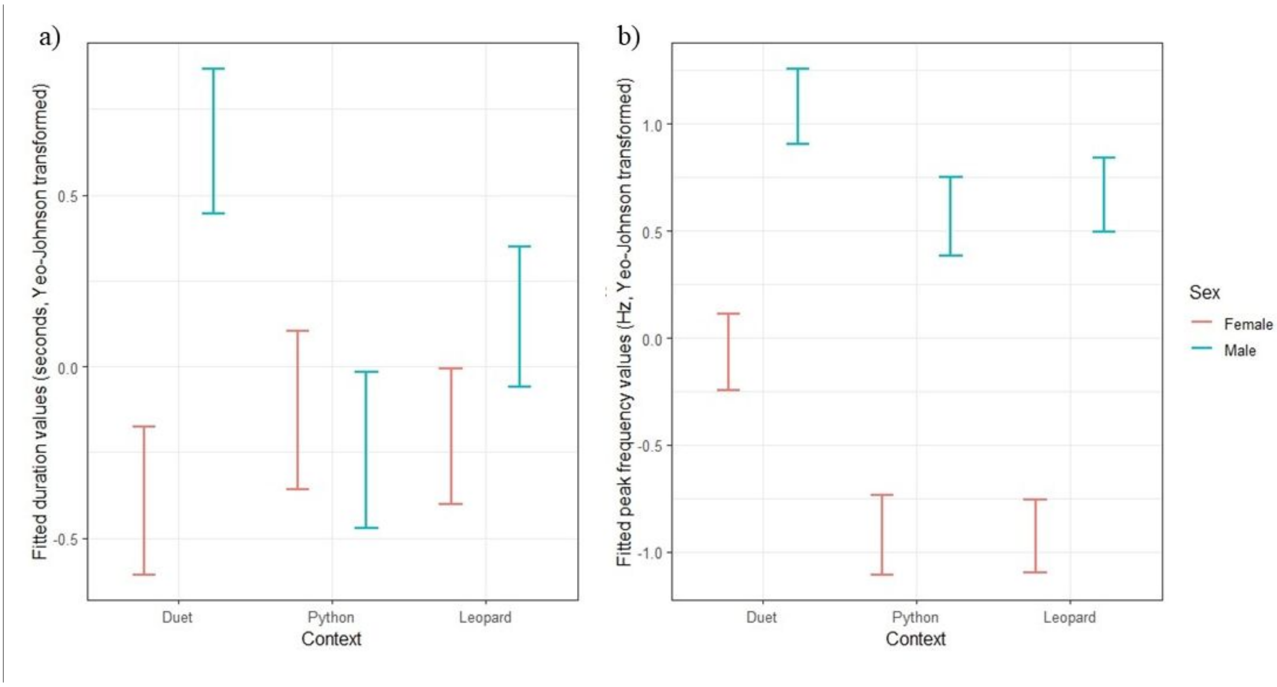

Figure 4.1. Effect plots of significant two-way interactions between sex and context for a) 'hoo' note duration and b) peak frequency. Bars represent fitted value estimates with 95% confidence intervals.

For low frequency, the full model differed significantly from its null model (χ^2(5 =211.41, $P<0.001$) with no significant interaction between context and sex ($\chi^2(2)=4.844$, $P=0.089$). Once the interaction term was removed, we found that low frequencies were affected separately by context (LRT, $\chi^2(2)=188.60$, $P<0.001$) and sex (LRT, $\chi^2(1)=20.06$, $P<0.001$). Females, again, were lower-pitched than males (Table S4.4) with both nevertheless uttering lower 'hoo' notes in python and leopard compared to duet songs, with no difference between predator songs (duet: 466.2 ± 52.3 Hz, python: 387.4 ± 61.0 Hz, leopard= 390.0 ± 50.8 Hz; Table S4.4).

For delta frequency the full model differed from its null model ($\chi^2(5)=51.68$, $P<0.001$) again with no interaction between context and sex ($\chi^2(2)=0.81$, $P=0.668$). Once the interaction term was removed, delta frequency was affected separately by context (LRT, $\chi^2(2)=43.78$, $P<0.001$) and sex (LRT, $\chi^2(1)=9.25$, $P<0.01$). Males had larger delta frequencies than females (Table S4.4), and both nevertheless produced ‘hoo’ notes with shorter delta frequencies in duet than python and leopard songs, with no differences between predator songs (duet: 83.1 ± 25.7 Hz, python: 113.3 ± 38.8 Hz, leopard= 109.9 ± 37.4 Hz; Table S4.4).

Finally, for ‘hoo’ inter-call intervals the full model differed significantly from its null model ($\chi^2(5) = 53.55$, $P<0.001$) with no significant interaction between context and sex ($\chi^2(2)=3.65$, $P=0.161$). Once the interaction term was removed , inter-call intervals were affected by context (LRT, $\chi^2(2)=49.336$, $P<0.001$), but not sex (LRT, $\chi^2(1)=0.62$, $P=0.431$): Intervals were longer in duet than leopard and python songs, with shortest interval in python songs) (Figure 4.2, Table S4.4).

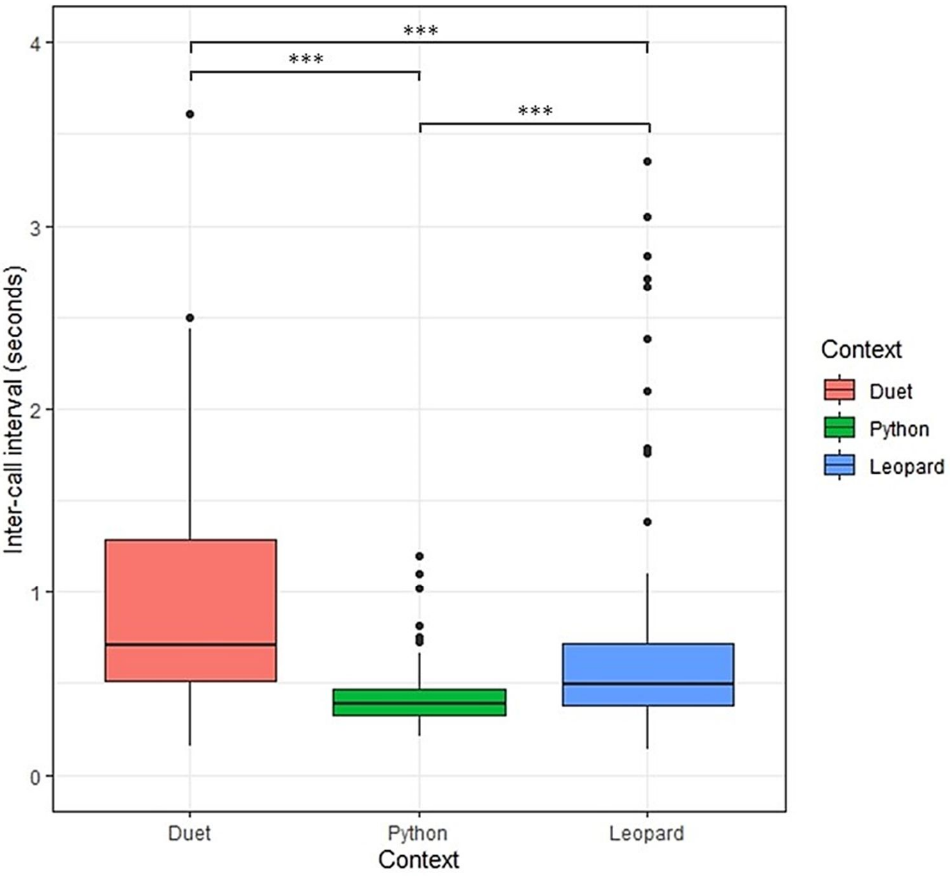

Figure 4.2. Box plots of the inter-call intervals duration between ‘hoo’ notes (untransformed data) across singing contexts (accounting for both males and females). Box indicates the first and third interquartile (25 and 75% of the data), black line shows the median, whiskers extend to the lowest and highest values within 1.5 times the inter-quartile range (IQR). Outliers are represented by black dots. (‘***’ $P < 0.001$).

Phonological structure of songs

Duration of introductory sequences (i.e. latency to first great call) and entire songs differed significantly across contexts (Figure 4.3, Table 4.2). Introductory sequences were significantly longer in leopard and python compared to duet songs (*P*<0.001 for both), with no differences between the two predator songs (*P*=0.541). Python and leopard songs were significantly longer than duet songs (*P*<0.001 for both), while leopard songs were significantly longer than python songs (*P*<0.01, see Table S4.5 for detailed pairwise comparisons).

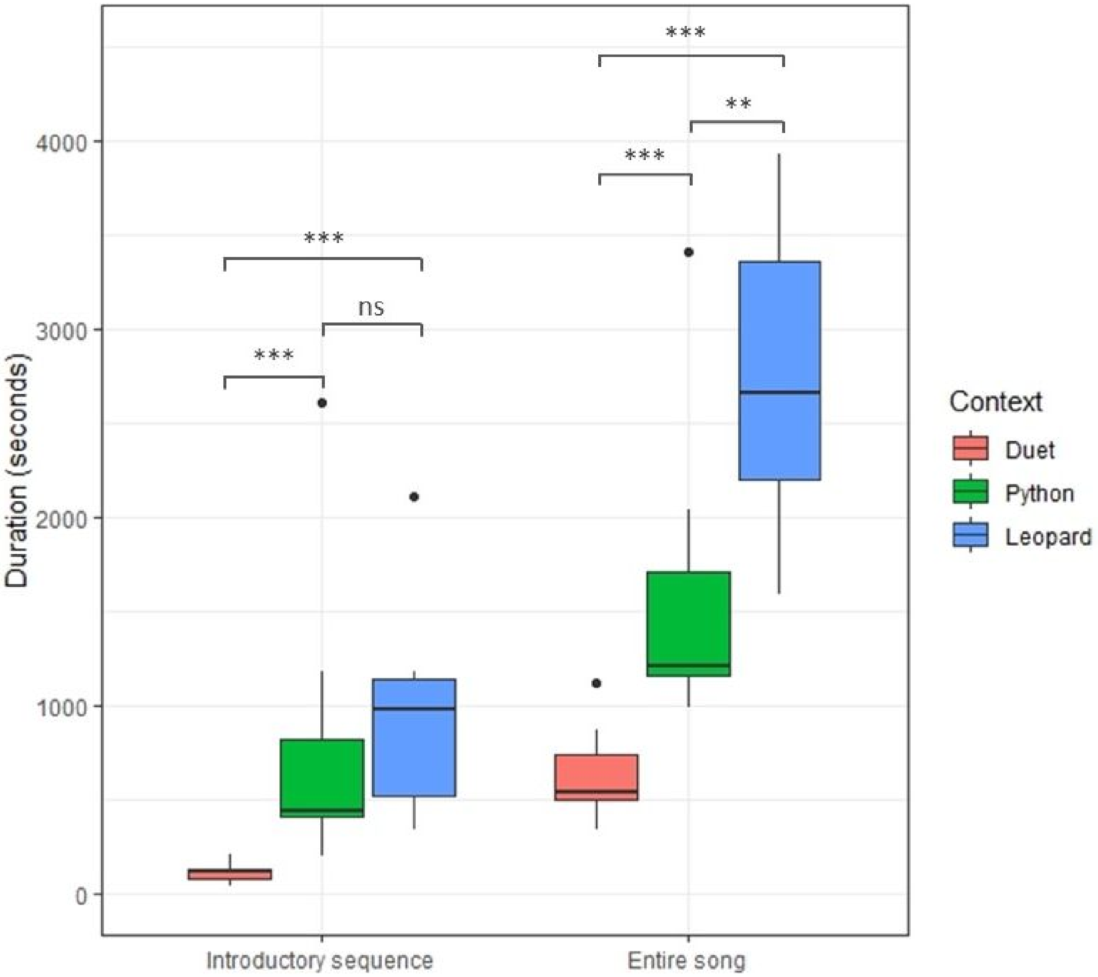

Figure 4.3. Box plots of duration of introductory sequence and total song duration across contexts (untransformed data). Box represent IQR (25% and 75%), black line shows the median, whiskers extend to the lowest and highest values within 1.5 times the IQR. Outliers are represented by black dots. (‘*’ $P < 0.05$, ‘**’ $P < 0.01$, ‘***’ $P < 0.001$, ‘ns’ non-significant).

Table 4.2. **Comparison of acoustic features of duet and predator songs** (linear mixed models, N_{duets}=11, N_{python}=7, $N_{leopard}$=11).

Acoustic parameters	Singing context					
	Duet mean ± SD	Python mean ± SD	Leopard mean ± SD	df	LRT	*P*
Introductory sequence (s)$^{\#}$	107.7 ± 50.6	813.7 ± 851.2	923.8 ± 532.2	2	39.34	< 0.001*
Song duration (s)	615.1 ± 232.5	1620.3 ± 857.7	2779.5 ± 808.0	2	45.98	< 0.001*
Hoo series (s)	11.5 ± 8.6	15.0 ± 12.05	25.3 ± 19.9	2	6.27	< 0.05*
Oo segment	3.6 ± 3.1	221.6 ± 89.2	59.9 ± 43.3	2	55.30	< 0.001*
Hoo-oo segment (s)	15.0 ± 8.8	236.6 ± 96.0	85.2 ± 45.5	2	50.62	< 0.001*

$^{\#}$ For analysis on introductory sequence duration, Group T leopard song was excluded from analysis as it contained no great call ($N_{leopard}$=10). (* P < 0.05).

We also found significant differences between duet and predator songs in the (regular) presence of ‘sharp wow’ notes, as a reminder we categorised each song as a function of ‘sharp wow’ notes being common (≥25) or rare (<25) (common; leopard: 11 of 11; python: 5 of 7; duet: 0 of 11; rare; leopard: 0 of 11; python: 2 of 7; duet: 11 of 11 – *P* < 0.001, N=29, Fisher’s Exact test – Leopard>Duet, *P* < 0.001; Python>Duet, *P* < 0.001; Python≈Leopard, *P*=0.412) and latencies to the first ‘sharp wow’ note (duet: 96.5 ± 52.9 s, N=11; python: 527.3 ± 429.1 s, N=7; leopard= 320.2 ± 137.3 s, N=11 – Kruskal-Wallis rank sum test, $\chi^2(2)$=18.72, *P*<0.001, N=28 – Python>Duet, *P*<0.001; Leopard>Duet, *P* < 0.001; Leopard≈Python, *P*=0.537).

The durations of the introductory ‘hoo’ note series, ‘oo’ note segment and ‘hoo-oo’ note segment also differed across contexts (Figure 4.4, Table 4.2). There were no differences in ‘hoo’ series duration across contexts (see Table S4.5 for detailed pairwise comparisons), but the subsequent ‘oo’ notes segment and the total ‘hoo’ and ‘oo’ notes segments (i.e. ‘hoo-oo’ segment) were both shorter in duet than predator songs, with longer durations for python than leopard songs (Table S4.5).

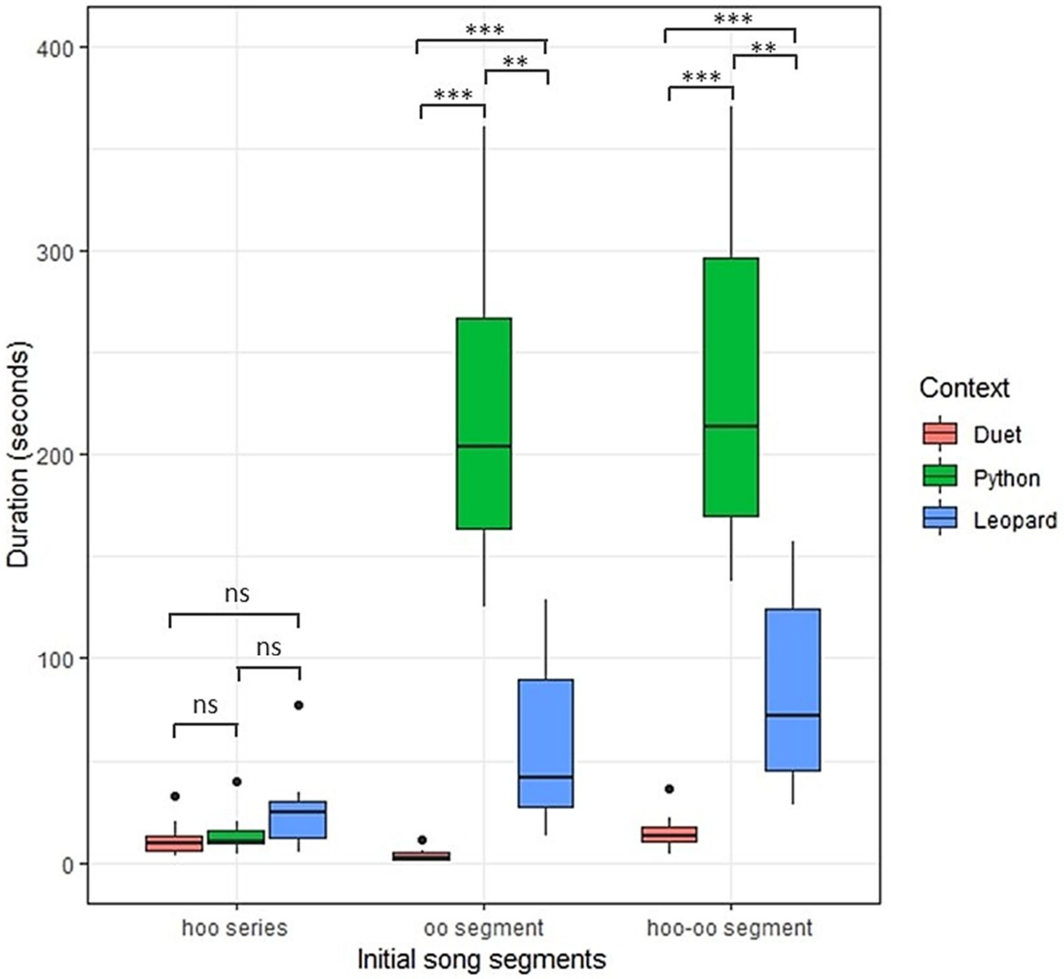

Figure 4.4. Box plots of the duration of the initial song segments across singing contexts (untransformed data). Box represent IQR (25% and 75%), black line shows the median, whiskers extend to the lowest and highest values within 1.5 times the IQR. Outliers are represented by black dots. ('*' $P < 0.05$, '**' $P < 0.01$, '***' $P < 0.001$, 'ns' non-significant).

Hoo-oo figures were present in the initial 'hoo-oo' notes segment in 6 of 7 python, 7 of 11 leopard and 0 of 11 duet songs. Python songs were found to contain more *Hoo-oo figures* in the initial segment than leopard songs (python: N=103, leopard: N=55, Comparison of Poisson rates, P<0.001), with longer inhaling parts in python than leopard songs (mean duration, python: 0.39 ± 0.14 s, leopard: 0.23 ± 0.13 s – U= 4576.5, P<0.001, Mann-Whitney test, Figure 4.5).

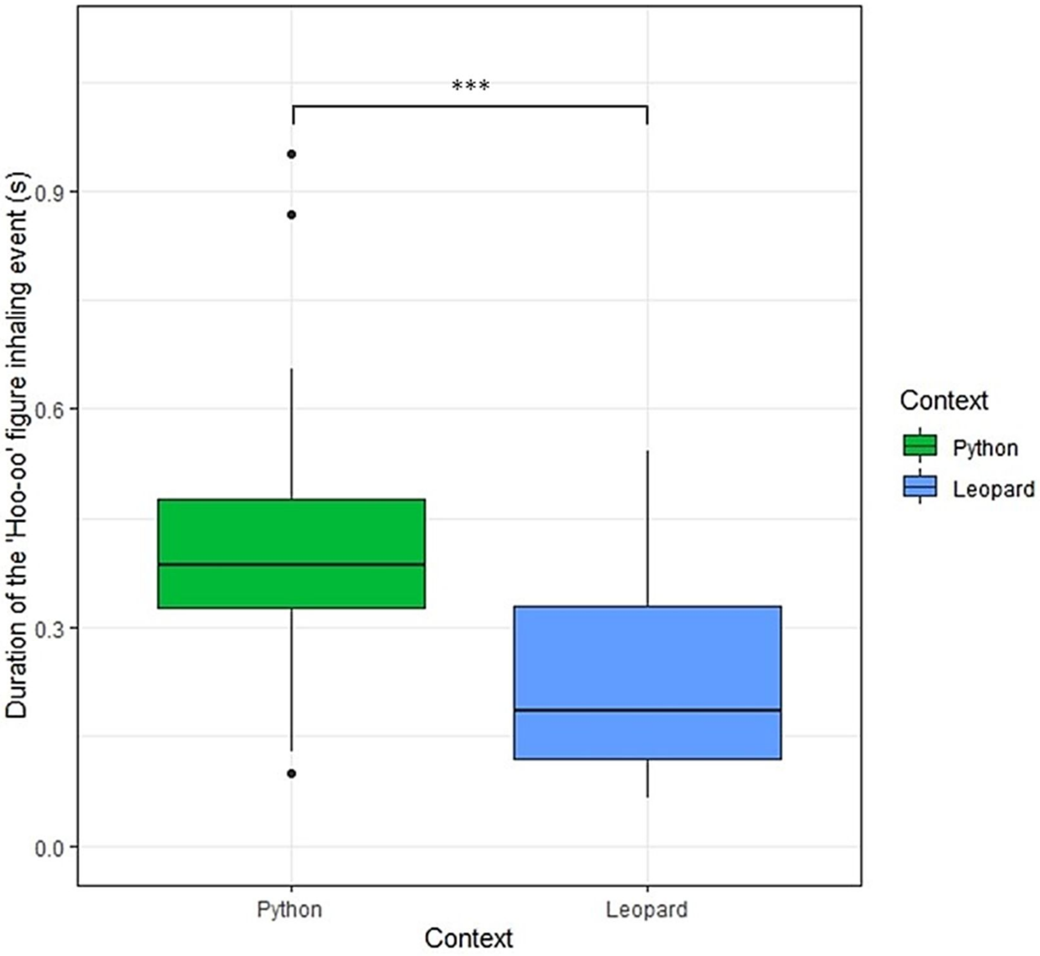

Figure 4.5. Box plots of the duration of the inhaling event produced in *Hoo-oo figures* between predator songs (N_{python}=103, $N_{leopard}$=55). Box represent IQR (25% and 75%), black line shows the median, whiskers extend to the lowest and highest values within 1.5 times the IQR. Outliers are represented by black dots. ('***' $P < 0.001$).

In the great call sequences, male replies to female great calls varied significantly between contexts (Kruskal-Wallis rank sum test, $\chi^2(2)=46.98$, $P<0.001$, N=106, Figure 4.6). They overlapped with their female great calls in 81.6% of leopard songs (N=38 sequences), 36.4% of python songs (N=22 sequences) and only 17.4 % of duet songs (N=46 sequences). The intervals between great call and male replies were shorter in leopard than python and duet songs ($P<0.001$ for both comparisons), and shorter in python than duet songs ($P<0.05$). Male replies differed across contexts: coda responses were present in 100% of duet songs (N=46 sequences), 5.3% of leopard songs (N=38 sequences), and 72.7% of python songs (N=22 sequences). In the remaining cases, male replies were singing follow-ups consisting of mainly 'sharp wow' notes.

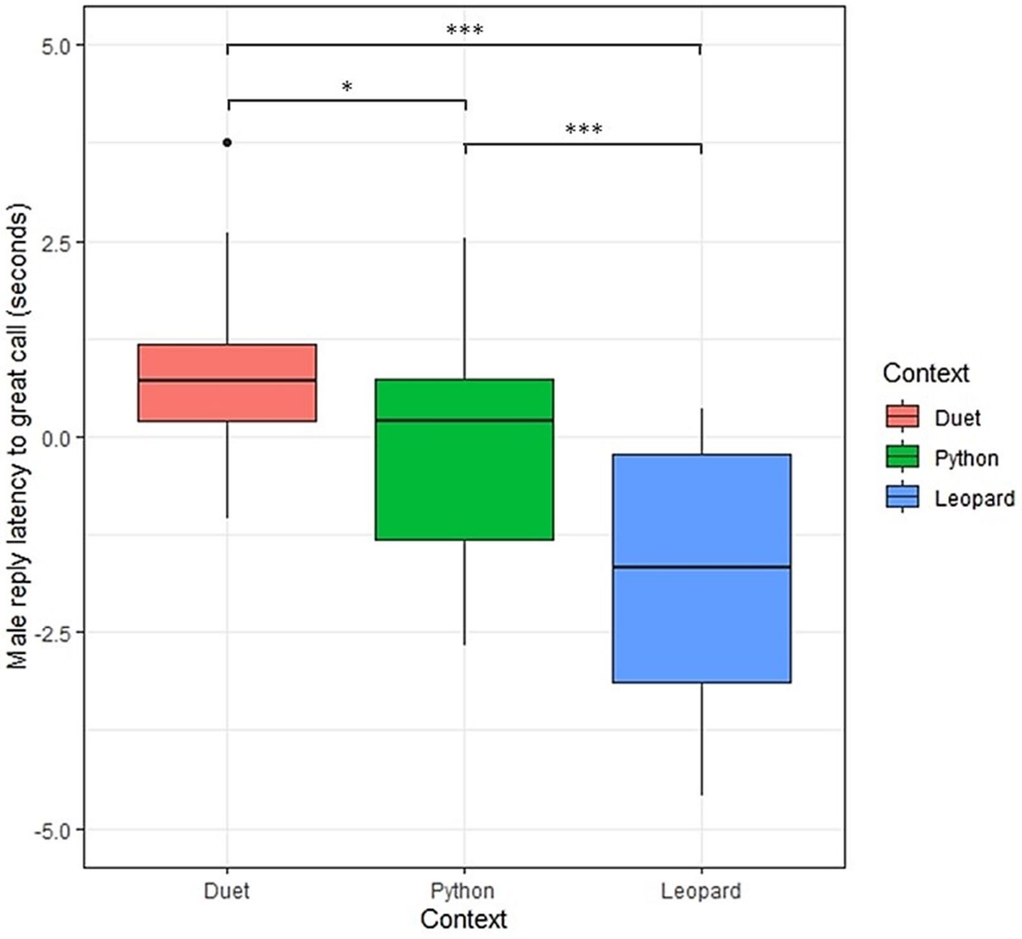

Figure 4.6. Latency of male replies to female great calls across contexts. Box plots represent IQR (25% and 75%), black line shows the median, whiskers extend to the lowest and highest values within 1.5 times the IQR. Outliers are represented by black dots. (‘*’ *P* < 0.05, ‘**’ *P* < 0.01, ‘***’ *P* < 0.001, ‘ns’ non-significant).

4.5 Discussion

White-handed gibbon song compositions differ between non-predatory and predatory contexts (Clarke et al., 2006) but whether songs can refer to different predator types is currently unknown. In this study, we carried out systematic comparisons of gibbon songs to clouded leopard and reticulated python models. Similar to what has already been described for clouded leopards (Clarke et al., 2006), subjects responded to reticulated python model by rapidly descending in canopy height, increasing their vigilance, and monitoring the model, matching the behaviour observed during natural snake encounters (Neudenberger, 1993; Uhde and Sommer, 2002; supplementary section). Regarding vocal responses, the python model triggered singing in 10 of 11 groups, mostly full predator songs (7 of 10 groups) and rarely as ‘soft hoo-oo songs’ (3 of 10 groups). Soft songs have previously been reported by Clarke et al (2006) using an apparently less stimulating python model .

In our study, the vocal responses to the python and leopard models showed all key indicators of predator songs, i.e. long duration, delayed first great call, rapid and frequently overlapping male replies and domination of ‘sharp wow’ notes. In contrast, the duet songs of 9 of 11 groups only contained few ‘sharp wow’ notes, all of which were given during duet songs that overlapped with those of neighbouring groups. The remaining two groups produced duet songs without sharp wows, suggesting that there might be two duet song subtypes, a topic for future research. Interestingly, occasional sharp wows have also been reported in non-habituated groups singing to human intruders (Clarke et al., 2006), but also in spontaneous duets or group duet responses to playback experiments broadcasted at closed distance (Andrieu et al., 2020) suggesting that this note type is linked to perceived aversiveness (Naguib and Mennill, 2010) or a social threat (i.e. territorial or mate defense) (Andrieu et al., 2020).

The main finding of our study concerns the structural differences in songs to leopard and python models, in comparison to non-predatory duet songs. Overall, python-elicited songs qualified as predator songs, although in some acoustic parameters they were intermediate between duet and leopard songs, notably regarding overall duration and male replies to female great calls. What could account for the consistent structural differences across the predator songs? One possibility is that leopards are perceived as more dangerous, due to their more versatile hunting behaviour, which may impact on individuals’ psychological states and make coordination of singing behaviour more difficult. African leopards have been shown to abandon hunting attempts in response to primate alarm calling (Zuberbühler et al., 1999) suggesting that gibbons may implement the same strategy when increasing song duration, while warning genetically related neighbours at the same time (Matsudaira et al., 2018; Andrieu et al. submitted a).

Gibbon songs consistently start with soft ‘hoo’ notes followed by the production of ‘oo’ notes interspersed by ‘hoo’ notes, before producing more complex notes assemblages. We were interested in these initial song segments, mainly because information about external events should be conveyed as early as possible. Clarke et al. (2015) already found context-dependant differences in some acoustic features of ‘hoo’ calls, notably differences in low and delta frequency features and inter-call intervals. We confirmed this pattern by showing that hoos in duet and predator songs differed in frequency-related parameters although we found no distinctions between predator types. Although white-handed gibbons do not produce full predator songs to raptors, they can respond with series of ‘hoo’ notes (Clarke et al., 2015), suggesting that comparisons of initial ‘hoo’ series should include responses to raptors. Clarke et al (2015) reported acoustically distinct ‘hoo’ calls to raptors encountered in the canopy, suggesting that the acoustic structure of ‘hoo’ calls may varies with predator location.

Nevertheless, we found predator-specific features in the early soft ‘hoo’ series, but in their delivery call, such as found in Campbell’s monkeys (Lemasson et al., 2010). Inter-call intervals between ‘hoo’ notes were longest in duet songs, intermediate in leopard songs and shortest in python songs. Adding Clarke et al’s (2015) raptor ‘hoos’ to the comparison, inter-call intervals were even longer in raptor than in duet songs, suggesting that group mates could base their decision making on this simple feature, although actual differences were very small (in the range a few tenths of seconds). Other context-specific structural features of the early introductory segments concerned the presence of *Hoo-oo figure*. These were absent in duet songs, whereas they were more frequent and acoustically different in python compared to leopard songs with longer inhaling events in the python context, and again this acoustic difference that might help recipients to discriminate between predator types.

In response to natural or experimental python encounters, gibbons can either produce ‘soft Hoo-oo songs’ (only soft ‘hoo’ and ‘oo’ notes, identical in note structure to the early ‘hoo-oo’ notes segment) or full and loud predator songs introduced by a longer ‘hoo-oo’ note segment compared to duet and leopard songs. Gibbons might be able to further develop this introductory ‘hoo-oo’ notes segment into full and loud predator songs, suggesting that these ‘soft Hoo-oo songs’ could be undeveloped predator songs; or that gibbons naturally produce two sub-types of predator songs. But still, it remains unclear how and why gibbons can produce these two types of singing responses when confronted, naturally or experimentally, to pythons.

3.6 Conclusion

We found consistent differences in white-handed gibbon songs given to two of their main predators, clouded leopards and reticulated pythons. These differences were already present in the early parts of the song, mainly in terms of call rates (i.e. early soft ‘hoo’ notes) and combinatorial meaning-bearing units (i.e. notes: ‘Sharp wow’; figures : ‘Hoo-oo’ figures; phrases: structure of great call sequences; segments: initial ‘Hoo-oo’ segment). In a companion study, we demonstrated that white-handed gibbons have the ability to discriminate duet from leopard songs (Andrieu et al., 2020); our current findings suggest it might also be the case across the predator songs. Although there is good evidence for predator-specific (or context-specific) alarm calls in many groups of animals, current evidence is usually in terms of differences in acoustic structure of individual call types. Here, we show that predator alarm calling can also be based on meaning-bearing unit differences other than syntax or following permutational rules (Zuberbühler, 2019b), due to differences in call rates, use of call notes repertoire or meaning-bearing units. Further research should focus on how other gibbon species respond to their natural predators, in order to test hypotheses about the impact of predation pressure on the ontogeny and evolution of song complexity (Stephan and Zuberbühler, 2008, 2014). Finally, further studies should investigate whether gibbons encode the location of danger, i.e., elevation and proximity, as has been demonstrated in titi monkeys and other primates (Zuberbühler, 2000b; Berthet et al., 2019).

4.7 Acknowledgments

We are thankful to the National Research Council of Thailand (NRCT), the Department of National Parks, Wildlife and Plant Conservation (DNP), and the superintendent of Khao Yai National Park for giving us the permission to conduct this research. We would like to especially thank Cathy Cebria and Françoise Cebria for having, based on paintings of a real snake, manufactured the reticulated python model; without their precious help, this study would not have been possible. We also thank Melanie Jackson, Prayoon Saenkhot, Attaklab Chaiyawat, and Surasak Homros for their help in the field. The research has been funded by Leverhulme Trust (Research Leadership Award F/00268/AP) and the European Research Council (grant number FP7; PRILANG GA283871).

4.8 Supplementary information

Figure S4.1. Models used in the experiments. Reticulated python model (left) and clouded leopard model (right).

Table S4.1 Group compositions of the white-handed gibbon population at the Mo Singto-Klong E-Tau study site (August 2014)

Group	N individuals	Identities and age/sex class
A*	3	1AM, 1AF
B*	5	2AM, 1AF, 1JF, 1I?
C	3	1AM, 1 AF, 1I?
E	3	1AM, 1AF, 1JM?
H	4	1AM, 1AF, 1JF, 1I?
M	5	1AM, 1AF, 1SAF, 1JM, 1I?
N*	6	2AM, 1AF, 1SAM, 1JF, 1I?
R	4	1AM, 1AF, 1AF, 1I?
S	3	1AM, 1AF, 1JM
T	5	1AM, 1AF, 1SAM, 1JM, 1I?
W*	4	2AM, 1AF, 1I?

*Multi-male group, primary male in bold; M=male, F=female, ?=sex unknown, A=Adult (age >8years), SA= Sub-adult (5-8 years), J=juvenile (2-5 years), I=infant (<2years).

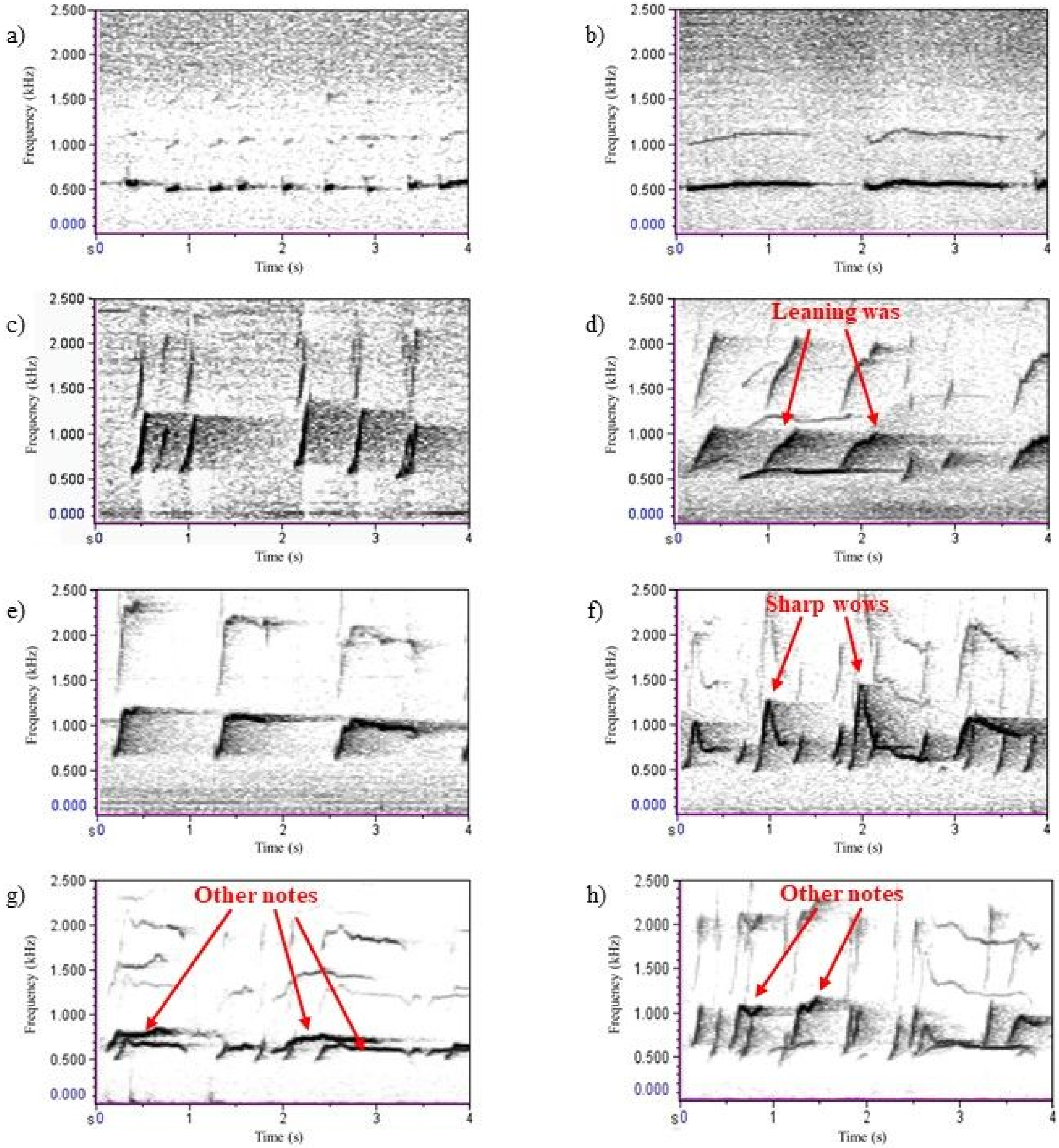

Figure S4.2. Song notes repertoire of white-handed gibbons (Raemaekers et al., 1984a; Clarke et al., 2006)**:** a) ‘hoo’ note; b) ‘oo’; c) ‘wa’; d) ‘leaning wa’; e) ‘wa-oo’; f) ‘sharp wow’; g) and h) ‘other notes’. Songs were digitized using Cool Edit Pro 2.1. spectrograms were drawn using 21.6 Hz filter bandwidth, 2.69 Hz frequency resolution, 33.3 ms time grid resolution and a Hanning window function.

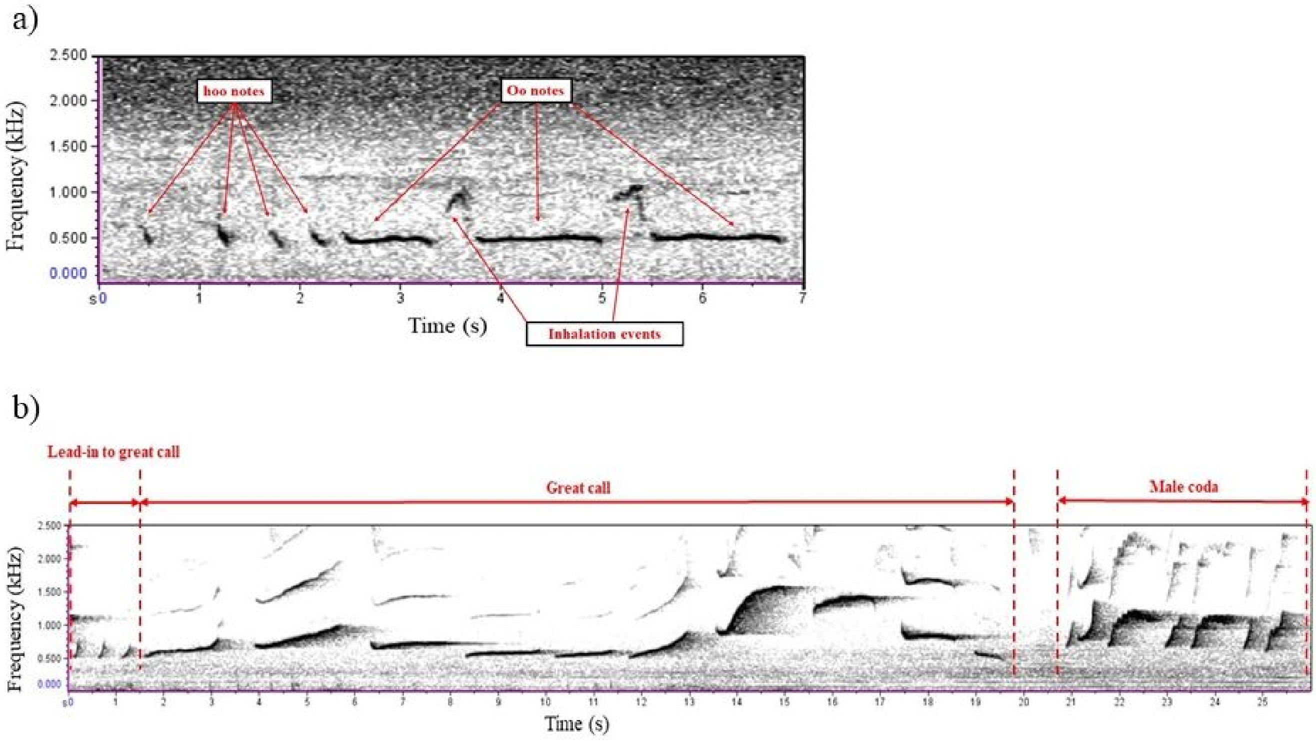

Figure S4.3. Spectrograms of a) an *Hoo-oo figure* and b) great call sequence: the female's great call and its mate's coda. In **a)** the exhaling notes are 'oo' notes separated by an inhaling event. *Hoo-oo figures* are mostly found in initial song segments but can appear later throughout the song bout. In **b)** the female great call is introduced by the 'lead-in to great call' phase produced by the female followed by the great call sequence (i.e. the female great call and the male coda). Spectrograms were set for fine temporal resolution (Hanning window function; 256-point fast Fourier transformation, 21.6 Hz bandwidth, 2.69 Hz frequency resolution and 33.3 ms grid time resolution).

Natural predator encounters

We witnessed two encounters between gibbons and reticulated pythons. The first one (group M) took place on 20 April 2014 at 0827h, local time (101°22'10.48" E, 14°26'29.43" N) with the adult male jumping into an adjacent feeding tree in which an approximately 5-metre-long reticulated python was located, about 10 m away from the group. Upon discovery, the male immediately started to utter a 'hoo' notes series. The female, located 20 m behind the male, also began producing a 'hoo' notes series, 64 seconds after the male 'hoo' series onset.

At 0828h, the male defecated and started vigorously moving branches of the python's adjacent tree, branches that were entangled to those supporting the python. Immediately after (0828h), all group members (the mated pair, one juvenile male, and one juvenile female) produced ‘oo’ notes interspaced with soft ‘hoo’ notes. At 0830h, a variable squirrel (*Callosciurus finlaysoni*) started to utter alarm calls. The juveniles never approached the python but stayed in the area, in a tree located approximately 8-10 m away during the whole predator encounter duration. At 0832h, the male shook the branches a second time whereas both females defecated in turn, while continuously singing ‘hoo’ and ‘oo’ notes, with the mated pair alternatively approaching and moving away from the adjacent feeding tree where the python was located to the tree where juveniles were.

At 0834h, the python started to move away, slowly pursued by the female at a distance of about 2 meters. At 0835h, the female approached the tree with the moving python in an apparent but unsuccessful attempt to grab the snake’s tail. The female then grabbed and shook a nearby liana, causing the python to fall 2m down into the lower parts of the tree (Figure S4.4). The female then moved away, joining back the tree where the mate and the juveniles were, and they all began to scan the ground and surrounding trees in an apparent attempt to locate the snake which was no longer visible. At 0841h the male dropped branches. At 0840h, the group finally left the area and also stopped vocalising, scanning the environment, and then left the site producing sporadic ‘hoo’ notes for two additional minutes. The group produced a ‘soft Hoo-oo song’ in response to their encounter with the reticulated python for a total duration of 13 min and 44 seconds. Their song was only composed of ‘hoo’ and ‘oo’ notes, interspersed by five *Hoo-oo figures*.

Figure S4.4. Gibbon’s attempt to make a reticulated python fall from trees.

The second python encounter happened on 8 July 2014 at 0826h, local time (101°22’27.44” E, 14°26’.14.11” N) between group S (mated pair, one adolescent male) and a python measuring about 2 m traveling at 15 m height. All group members mobbed the python, with the mated pair frantically shaking branches for about 10 s. Again, this caused the python to fall and disappear in the undergrowth.

During branch-shaking, the pair initially produced ‘hoo’ and ‘oo’ notes series that ended up by an *Ooaa burst figure*. After the python had fallen, the group immediately moved approximately 15m away, carrying on with singing and vigilance. After a few minutes, the group continued to move further away but started to sing a long and loud predator song for a total duration of 34 min after which they were about 140m away from where they had found the python. The final song consisted of 19 *Hoo-Oo figures*, 18 *Ooaa burst figures* and the production of 8 great call sequences; this triggered an Ooaa duet in an adjacent group.

Table S4.2 Behavioural variables used to assess anti-predator behaviours during predator model presentation

➢ **Behavioural activity**	
Feeding	Handling or consuming food items
Resting	Stationary position, standing still with or without eyes closed
Grooming	Autogrooming, allogrooming (giver and receiver identity were collected)
Social	Sexual behaviour, play, aggression, and parental care
Moving	Travel within or between trees (at least 2 metres)
Vigilance	Scanning the environment, head rotating by at least 45° **(Koenig, 1998)**
Other	Behaviour not classified into any of the above categories
➢ **Body Position**	
Hanging	Suspended in the air, grabbing a branch or a tree part with at least one arm
Sitting/Lying down	Sitting on a branch or on the ground / Resting in horizontal position
➢ **Gaze direction (staring at a specific location/direction/animal/person for ≥ 3s)**	
Predator model/Ground	Staring directly at the predator model / Looking towards or scanning actively the ground
Canopy	Looking around, or towards a specific location in the trees at the same elevation as the animal location
Sky	Looking up at the sky
Group member	Looking at a group member (the identity of the receiver was collected)
Observer	Looking at the observer
Elsewhere	Looking in a direction that cannot be classified into any of the above category
Nowhere	Resting with eyes closed
➢ **Other measurements**	
Elevation (m)	Height of the animal in relation to the ground
Model inspection (m)	Distance of individuals in relation to the model
Proximity (m)	Distance between the 2 focal individuals (paired male and female)
Defecation/Urination	Depositing faeces and/or urine
Dropping branches	Individuals shaking branch(es) from the tree onto the ground

Table S4.3 Anti-predator behaviour of white-handed gibbons to python model (N=11 model presentations, with a total of 376 scan sampling observations for 22 individuals, Exact Wilcoxon signed-rank Permutation test with 1000 iterations).

Variable		Detection		Statistics	
		Before Mean ± SD	After Mean ± SD	Corrected V value[†]	Corrected P value[†]
Body position	Hanging	4.4 ± 2.4	5.9 ± 2.5	7.9	0.224
	Sitting/lying	4.1 ± 3.8	2.7 ± 2.6	30.9	0.262
Activity	Moving	1.5± 1.1	1.2 ± 1.1	30.3	0.572
	Feeding	3.7 ± 2.6	0.2 ± 0.5	60.7	< 0.01*
	Resting	2.4 ± 1.9	0.1 ± 0.5	55.0	< 0.01*
	Grooming	0.8 ± 0.3	0.7 ± 1.1	14.3	0.730
Gaze	Canopy	7.6 ± 2.7	2.0 ± 2.2	60.7	< 0.01*
	Group member	0.6 ± 0.9	0.5 ± 1.0	9.4	0.678

Mean behaviours occurrence was calculated over 376 scan sampling, i.e. 188 observations per condition (i.e. before and after). 'Social' and 'Other' activity categories, as well as gazes towards the 'sky', 'observer' and qualified as 'nowhere' were excluded from the analysis due to their respective total small number of observations (4, 1, 3, 2, 2 respectively).

[†] Sample size reduced to 188 scan sampling observations for 11 individuals when correction was applied. (* $P < 0.05$).

Table S4.4 Pairwise comparisons of frequency-related acoustic parameters of the 'hoo' notes emitted by males and females across singing contexts (Difference of Least Square means, with Bonferroni corrections).

Context x Sex 1	**Context x Sex 2**	Estimate	Std. Error	df	t.ratio	p
Duration						
Duet x Female	Python x Female	-0.266	0.216	207.4	-1.229	1
	Leopard x Female	-0.189	0.184	207.1	-1.025	1
	Duet x Male	-1.049	0.315	45.5	-3.334	< 0.05*
	Python x Male	-0.149	0.326	50.3	-0.456	1
	Leopard x Male	-0.537	0.309	42.2	-1.736	1
Python x Female	Leopard x Female	0.077	0.200	209.2	0.386	1
	Duet x Male	-0.784	0.326	51.4	-2.401	0.3
	Python x Male	0.117	0.337	56.2	0.347	1
	Leopard x Male	-0.271	0.321	48.0	-0.845	1
Leopard x Female	Duet x Male	-0.861	0.301	38.6	-2.856	0.103
	Python x Male	0.040	0.313	43.3	0.127	1
	Leopard x Male	-0.349	0.296	35.5	-1.179	1
Duet x Male	Python x Male	0.901	0.204	209.1	4.420	< 0.001*
	Leopard x Male	0.512	0.189	214.1	2.716	0.107
Python x Male	Leopard x Male	-0.389	0.202	215.1	-1.920	0.843
Peak frequency						
Duet x Female	Python x Female	0.856	0.133	202.9	6.415	< 0.001*
	Leopard x Female	0.861	0.113	202.9	7.592	< 0.001*
	Duet x Male	-1.150	0.263	34.1	-4.370	< 0.01*
	Python x Male	-0.636	0.268	36.9	-2.368	0.349
	Leopard x Male	-0.734	0.261	32.8	-2.816	0.122
Python x Female	Leopard x Female	0.005	0.124	203.9	0.044	1
	Duet x Male	-2.005	0.268	37.0	-7.469	< 0.001*
	Python x Male	-1.492	0.274	39.8	-5.446	< 0.001*
	Leopard x Male	-1.590	0.266	35.7	-5.973	< 0.001*
Leopard x Female	Duet x Male	-2.011	0.257	31.0	-7.827	< 0.001*
	Python x Male	-1.497	0.263	33.7	-5.703	< 0.001*
	Leopard x Male	-1.595	0.254	29.8	-6.269	< 0.001*
Duet x Male	Python x Male	0.514	0.126	204.3	4.077	< 0.01*
	Leopard x Male	0.415	0.117	208.0	3.547	< 0.01*
Python x Male	Leopard x Male	-0.098	0.126	207.6	-0.782	1
Condition 1	**Condition 2**	Estimate	Std. Error	df	t.ratio	p
Low Frequency						
Duet	Python	1.100	0.077	202	14.349	< 0.001*
	Leopard	1.070	0.069	204	15.701	< 0.001*
Python	Leopard	-0.031	0.074	204	-0.416	1
Female	Male	-1.05	0.192	24.2	-5.463	< 0.001*
Delta frequency						
Duet	Python	-0.745	0.141	206	-5.287	< 0.001*
	Leopard	-0.808	0.125	208	-6.467	< 0.001*
Python	Leopard	-0.063	0.135	210	-0.468	1
Female	Male	-0.829	0.258	24.6	-3.216	< 0.01*

Table S4.5. Pairwise comparisons across singing contexts in terms of song structure (Difference of Least Square means, with Bonferroni corrections).

Context 1	**Context 2**	Estimate	Std. Error	df	t.ratio	p
Introductory sequence						
Duet	Python	-1.84	0.243	20.6	-7.565	< 0.001*
	Leopard	-2.19	0.212	19.5	-10.326	< 0.001*
Python	Leopard	-0.349	0.252	21.2	-1.385	0.541
Song duration						
Duet	Python	-0.938	0.177	23.8	-5.293	< 0.001*
	Leopard	-1.530	0.153	20.3	-9.985	< 0.001*
Python	Leopard	-0592	0.177	23.8	-3.340	< 0.01*
Hoo series						
Duet	Python	-0.217	0.335	22.7	-0.648	1
	Leopard	-0.728	0.285	20.2	-2.553	0.057
Python	Leopard	-0.511	0.335	22.7	-1.527	0.421
Oo segment						
Duet	Python	-4.43	0.391	23.8	-11.314	< 0.001*
	Leopard	-2.93	0.338	20.3	-8.662	< 0.001*
Python	Leopard	1.50	0.391	23.8	3.825	< 0.01*
Hoo-oo segment						
Duet	Python	-2.84	0.272	23.8	-10.459	< 0.001*
	Leopard	-1.74	0.235	20.3	-7.416	< 0.001*
Python	Leopard	1.10	0.272	23.8	4.047	< 0.01*

5- GENERAL DISCUSSION

5.1 Summary

The main objective of this thesis research was a detailed exploration of the mechanisms underlying the gibbon song communication system, which consisted of fieldwork on a wild population of white-handed gibbons, in Khao Yai National Park, Thailand. The first project aimed to investigate the mechanisms underlying natural duet singing interactions between groups. The main approach was to integrate existing demographic and genetic data with natural recordings of duet song interactions. The second project aimed to explore the abilities of gibbons to discriminate songs given in predatory and non-predatory situations and assign meaning to them. The main approach was to conduct playback experiments of different song types recorded in different contexts, i.e. duet and predator songs, and compare subjects' responses. The third project, finally, aimed to investigate whether gibbon song functioned as a referential alarm calling system. Here, the main approach was to present different predator models, recording subjects' responses to them and investigate the acoustic differences of songs given in response to different predator types.

In the first project, I built on previous observations showing that mated pairs of white-handed gibbons regularly produced morning duet songs that could transmit over long distances (Marshall and Marshall, 1976). Importantly, duetting is often contagious, especially between neighbouring groups, but during this process groups can either respect each other singing turn with turn-taking (by waiting until the previous group has finished with their song) or compete with counter-singing (by starting with their own song before a neighbouring group has finished with singing). Here, earlier observations showed that groups were more likely to respond with duet songs if the preceding songs were produced by neighbours, compared to more distant non-neighbouring groups (Raemaekers and Raemaekers, 1985b). Also, previous work suggested that song overlap was avoided in song exchanges between neighbouring groups, but not between non-neighbours (Raemaekers and Raemaekers, 1985b).

My own dataset contradicted some of these previous conclusions by showing that the default strategy in white-handed gibbons is competitive counter-singing. In some specific spatial, social and genetic conditions, however, individuals can switch from competitive counter-singing to turn-taking, which was the case in the following conditions. First, if two neighbouring groups consisted of single males who were related then this decreased the likelihood of counter-singing, especially at close proximity. Fathers and sons, brothers and half-brothers, in other words, appear to respect each other's duet song more than other pairs of males. Second, a similar kin effect was found between neighbouring groups consisting of (half-) brothers and sisters, but also fathers and daughters, and mothers and sons, but in this case, it was regardless of distance. Third, multi-male groups were always more likely to be targeted by counter-singing than single male groups, also regardless of distance. I then looked at the counterstrategies deployed by groups targeted by counter-singing and found that subjects responded by increasing the duration of their own songs. The counter-singing groups themselves, however, did not show any further adjustments. However, counter-singing groups were more likely to pause for one key component of gibbon song, the female's great calls, but only when interacting with a neighbouring group.

Overall, these results demonstrated that duet interactive singing in white-handed gibbons functions in sexual rivalry and intergroup spacing, with distance, group composition and kinship as mediating factors. These findings also provide the first evidence for kinship playing a role in vocally-mediated intergroup interactions.

The second project aimed to investigate the abilities of gibbons to discriminate and infer the meaning of different song types. Here, I only compared duet and leopard songs, due to time limits. While duet songs function in interactions with neighbouring groups, predator songs are given exclusively when facing a threat (Clarke et al., 2006; Raemaekers and Raemaekers, 1985b). Previous research has already shown that duet and predator songs are mostly identical in the note repertoire deployed by the callers, but also that there are consistent structural differences across contexts (Clarke et al., 2006, 2015). Duet songs can be heard over considerable distances, even beyond the emitters' territory, a capacity suggesting that these signals have evolved to communicate beyond the immediate social group (Raemaekers and Raemaekers, 1985b; Geissmann, 1999; Geissmann and Orgeldinger, 2000), especially considering that gibbon groups are often surrounded by neighbouring groups with genetic relatives (Matsudaira et al., 2018). The same could be the case for predator songs, suggesting that structural differences between duet and predator songs might be exploited as meaningful signals by out-group individuals. To test whether gibbons could discriminate and infer the meaning of different song types, I conducted a playback experiment that simulated natural singing events by neighbours (i.e. previously recorded spontaneous duet songs vs. leopard songs) to different target groups. As expected, groups reliably replied to 'simulated neighbours' duet songs playback by giving their own duet songs. When predator songs were broadcasted, however, tested groups remained mostly silent, although predator songs were sometimes heard from distant groups (in 2 of 6 trials). However, this cryptic vocal behaviour by the tested groups was supplemented by strong anti-predator behaviours typical for predator encounters (Clarke et al., 2012), such as an increased vigilance, monitoring actively the ground, and increased defecation rates, a sign of distress (Boissy, 1995), something that was never observed in duet playback trials.

From this set of evidence, I concluded that gibbons discriminated between the two song types, and thus could infer meaning from them, which confirms that gibbon song functions as semantic vehicles in long-distance communication.

In the third project, lastly, I investigated the function of gibbon song as a referential alarm calling system. To investigate this hypothesis, I fabricated a reticulated python model to overcome limitations of a previous study that failed to elicit systematic responses with this predator type (Clarke et al., 2006, 2012). I compared the responses to the new python model with responses to a clouded leopard model by recording songs elicited by the models. I then further compared the predator songs with natural duet songs to test whether gibbons arranged song units in specific ways that would allow others to infer meaning about the type of predator encountered by the caller. In my study, the python model reliably elicited strong singing and strong anti-predator responses, with decreased use of lower strata and increased gazing towards the predator model, combined with increased vigilance and defecation rates. Acoustic analyses of singing responses further revealed that both, leopard and python songs, qualified as predator songs (as opposed to duet songs), with longer overall song and introductory sequence durations, earlier male replies to female great calls, and regular production of 'sharp wow' notes in the later song bouts (Clarke et al., 2006). Also, several acoustic structures differed between predator and non-predator song types. In the early song bout segments, the delivery rate of the song opening 'hoo' notes, the duration of the following 'oo' segments and production of '*hoo-oo*' figures and its inhaling event duration, all differed between duet, python and leopard songs. Great call sequences also differed in predator compared to duet songs, with male 'coda' replies to female great calls generally being rare in leopard songs, erratic in python songs and mandatory in duets. Overall, I concluded that predator songs differed from each other and from non-predatory duet songs, potentially allowing conspecifics to infer meaning about the external events faced by the caller.

To sum-up, the three projects strengthen and provide new insights into the mechanisms underlying gibbon acoustic communication, including competitive intergroup interactions but also cooperative predator avoidance. This is achieved by the gibbons' ability to code songs in different structural arrangements allowing others to infer meaning, a basic requirement for referential communication.

5.2 Counter-singing in gibbons: implications for evolutionary theory

I investigated the spread of duet song production by two subsequent groups, by analysing previously ignored or unavailable group-specific and neighbourhood-specific variables, such as group composition and genetic relatedness between individuals. Duet song can spread in two ways, either by cooperative turn-taking or by competitive counter-singing.

Counter-singing is of ongoing interest and a major topic of debate among bird scientists (Naguib and Mennill, 2010; Searcy and Beecher, 2009, 2011). Birdsong is a classic example of animal long-distance signalling and, in many species, males overlap each other with counter-singing (Armstrong, 1973). Several hypotheses have been proposed for the function of this behaviour (reviewed in Helfer and Osiejuk, 2015). First, counter-singing could be a mere stochastic event, for example, due to limited processing abilities with no functional relevance (Searcy and Beecher, 2009, 2011). Second, overlap could function in a cooperative way, for example, if it leads to an amplified signal that benefits both individuals, the signal enhancement hypothesis (Narins, 1992; Greenfield, 1994b; Bates and Cropp, 2010; Rehberg-Besler et al., 2017). Third, overlap could function as an expression of agonistic motivation to signal an intent to escalate a competitive interaction, the aggressive signalling hypothesis (Kunc et al., 2006; Mennill and Ratcliffe, 2004a, 2004b; Naguib and Mennill, 2010). Fourth, and related to the previous, overlap could function as an expression of submission, to avoid forthcoming escalation, the submissive signalling hypothesis (Searcy and Beecher, 2009). Finally, overlap could function in mate attraction, such as an honest indicator of genetic quality or physical condition (Gil and Gahr, 2002).

Amongst the five hypotheses, the agonistic signalling hypothesis has received most support in bird studies. Also, in many species, victims of overlap adapt their behaviour by adjusting song length and singing rates or by approaching the opponent (Dabelsteen et al., 1997; Naguib, 1999; McGregor et al., 2000; Naguib and Todt, 1997). The agonistic signalling hypothesis further predicts that overlap should be particularly common in territorial species where it should occur as a function of proximity. As long-distance calls are subject to acoustic degradation with increasing distance, receivers can estimate signaller position and distance (Naguib and Wiley, 2001) and, if deemed necessary, engage in signal overlapping. For birdsong, Searcy and Beecher (2009, 2011) have argued that, across studies, overlap is less common than expected by chance, casting doubt on the agonistic signalling hypothesis. In contrast, Helfer and Osiejuk (2015), have argued that, across bird studies, victims of overlap usually respond by taking countermeasures, often accompanied by specific behaviour that suggests that counter-signing in birds is an aggressive signal (Naguib and Mennill 2010).

Within non-human primates, there is a lot of evidence for individuals responding to each other's calls, albeit mostly at close range and within groups, such as Diana monkey (*Cercopithecus diana*) or Campbell's monkey (*Cercopithecus campbelli*) contact calls (Uster and Zuberbühler, 2001), (Lemasson et al., 2011). Between groups, turn-taking has been reported in the long-distance songs of Siamangs (*Hylobates syndactylus*) (Geissmann and Orgeldinger, 1997), Mueller's gibbons (*Hylobates muelleri*) (Inoue et al., 2013) or white-handed gibbons (*Hylobates lar*) (Raemaekers and Raemaekers, 1985b). In some species, duet song spread is such that groups do not respect each other's acoustic signal, i.e. duetting becomes part of inter-group counter-singing, such as in Black-crested gibbons (*Nomascus concolor*), Kloss's gibbons (*Hylobates klossi*) and Bornean gibbons (*Hylobates muelleri*) (Tenaza, 1976; Mitani, 1985b; Fan et al., 2006). Here, songs produced interactively were longer than songs produce alone, suggesting that groups attend and respond to each other's songs.

Regarding the five hypotheses put forward for birdsong, the results of my first project demonstrated that counter-singing in white-handed gibbons is not a random by-product of chance (Searcy and Beecher, 2011), for example due to the fact that groups sing at different rates and do not anticipate the offset of a previous song (Naguib and Mennill, 2010). If overlap occurs due to such stochastic processes, then it is unlikely to be part of intergroup agonism or other sorts of social hostility (Searcy and Beecher, 2009, 2011). In my data, social, genetic and spatial factors all had an effect on duet song timing, due to the fact that subjects temporally respected or disrespected a neighbouring group's duet song, and due to counter-measures of groups that were targeted by counter-singing. My data are also difficult to reconcile with the signal enhancement hypothesis, which predicts that individuals amplify cooperatively their signal by singing together, for which there was no evidence in any way.

The remaining three hypothesis for song overlap, however, need to be considered more carefully for duet song interactions in gibbons. First, competitive counter-singing may function as a ritualised way of long-distance intergroup aggression, the aggressive signalling hypothesis, emanating from resource or territorial border defence. Under this hypothesis, counter-singing is part of an increasing scale of escalating a conflict. Although this hypothesis needs further testing, some of my findings are consistent. For example, the fact that song duration is increased by some initiator groups experiencing counter-singing may be a sign of a higher motivation to defend their resources and a will to escalate (or de-escalate) a conflict. Further evidence is that counter-singing rarely leads to physical aggression. Only 3 of 49 physical encounters preceded by counter-singing led to subsequent intergroup conflict.

Counter-singing is also in compliance with the mate attraction hypothesis, which states that song overlap could serve as an honest indicator of genetic quality or physical condition (Gil and Gahr, 2002). In my study, multi-male groups were more likely to elicit counter-singing than single male groups. Under this hypothesis, single male groups deploy counter-singing as a way to prevent future agonistic encounters and as an honest signal of competitiveness and motivation for mate guarding to prevent extra pair copulations (Barelli et al., 2013).

To date, the biological phenomenon of song overlap was studied systematically only in birds. My research has demonstrated that gibbon song provides a good model to complement the bird studies. Taking turns is easier to achieve when signals are short and soft, compared to the songs of birds. Gibbons may follow similar principles as birds when using their duet songs in interactions with neighbouring groups. My main conclusions that the so-called bird ‘song overlap’, is in white-handed gibbons, a signal and that it functions mostly in sexual rivalry and intergroup spacing.

5.3 Gibbon songs and human language

a. Social coordination in gibbon song

Human language is built on a signal coding system that allows signallers and receivers to interact in referential ways. Receivers can make complex inferences about the pertinence and meaning of a signaller's communicative behaviour, a process aided by a human ability to infer and share intentions between speakers and listeners (Tomasello et al., 2005; Townsend et al., 2017). But human language is also a highly interactive acoustic achievement by which speech signals are exchanged rapidly and with high level of precision and by which vocal overlap is avoided for a correct understanding and prevention of loss of information (Stokes and Williams, 1968).

This alternation of vocal interactions is termed turn-taking (Sacks et al., 1974). Bi-directional acoustic exchanges, or turn-taking, also exist in many animal species, requiring specific adaptations for both the signallers and the receivers. Yoshida and Okanoya, (2005) have classified turn-taking behaviour in animals as chorus, duet and antiphony. Choruses are defined as sex–specific vocal exchanges, usually by males, to attract mates, maintain social contact, and regulate intergroup spacing (Greenfield, 1994b). Duets, in contrast, are defined as overlapping bouts of vocalisations given by paired individuals, such that the elements within those bouts have a high level of alternation, or a low coefficient of variation of the intervals between their element or both, with a degree of temporal precision (Farabaugh 1982). Tropical wrens, belonging to the genus *Thryothorus*, represent a classic example of duetting birds that includes species producing from poorly synchronised to precise-timed coordinated *duets* (Farabaugh, 1985). Antiphony, finally, is defined as three and more animals uttering sounds in alternation to a preceding signal, a behaviour found mainly in environments where visual contact is difficult, requiring animals to remain in auditory contact (Yoshida and Okanoya, 2005). An example is the contact calls of many forest primates (e.g. Uster and Zuberbühler, 2001).

Many non-human primate species engage in turn-taking behaviour with conspecifics. In prosimians, antiphonal loud calls of pairs of *Lepilemur edwardsi* function in territorial defence, space regulation and cohesiveness between pairs and neighbours (Rasoloharijaona et al., 2006). In New World monkeys, marmosets (*Callithrix jacchus*) vocalisations are exchanged in alternation without overlapping the calls of conspecifics (Takahashi et al., 2013) and squirrel monkeys respect an order in vocalising depending on the recognition of other individuals' calls (Symmes and Biben, 1988). In Old World monkeys, the turn-taking rules were described for adult Campbell's monkeys (*Cercopithecus campbelli*) (Lemasson et al., 2011). In the Hominoïdae, vocal turn-taking has only been described for gibbons and humans, with data from Siamangs (*Hylobates syndactylus*) (Geissmann and Orgeldinger, 1997), Mueller's gibbons (*Hylobates muelleri*) (Inoue et al., 2013) and white-handed gibbons (Raemaekers and Raemaekers, 1985b).

As mentioned, vocal interactions can either be non-aggressive (i.e. turn-taking) or competitive (i.e. counter-singing) and exploring when and how animals opt for one or the other strategy could help explain how cooperative human turn-taking, i.e. speech, have evolutionarily emerged. Vocal turn-taking has not been described in great apes, although orang-utans (*Pongo sp.*), bonobos (*Pan paniscus*) and chimpanzees (*Pan troglodytes*) show turn-taking with gestural signals (Rossano and Liebal, 2014; Rossano, 2013; Call and Tomasello, 2007). Hylobatids are thus the only other member of the Hominoid group showing vocal turn-taking, with precise-timed signal alternations, observed not only in duet interactions between group, but also within duet song coordination within groups with synchronised vocal contributions between mated pairs. Vocal turn-taking seen in gibbons and humans is either the result of independent evolution in phylogenetically distinct groups (Von Helversen, 1980) or due to loss in great apes despite shared inheritance and ancestral emergence in the Hominoid lineage (Levinson, 2016). Future research will have to investigate these two hypotheses with targeted studies.

b. Reference in gibbon song

Conveying and decoding information is essential in human language and animal communication can elucidate the evolutionary path to language. Non-human primate studies have generated noteworthy progress in recent years, regarding the coding, inference and social intelligence needed for complex communication. Research in animal song has also been part of this investigation, often in terms of the combinatorial structure of different song types.

Nonetheless, how human language has evolved from primate communication is still largely unknown. The generally accepted conclusion is that some primate calls are carriers of distinct meanings and that this can been aided by an additional role of pragmatics (Seyfarth and Cheney, 2018), both in terms of allocating caller intention (Zuberbühler, 2000a) or general world knowledge (Arnold and Zuberbühler, 2013; Zuberbühler, 2000c).

For example, vervet monkeys (*Chlorocebus pygerythrus*) produce acoustically distinct calls to different predator types (Seyfarth et al., 1980a, 1980b), although the production of acoustically distinct calls is not the only one way by which primates communicate about external events. For example, Guereza colobus monkeys (*Colobus guereza*) specify the nature of a perceived danger by altering the number of calls per sequence (Schel et al., 2009) and other alarm call systems are based on graded changes of acoustic features or differences in call intensity (e.g. chacma baboons, *Papio cynocephalus ursinus*, Fischer et al., (2001a, 2001b). Finally, there is also evidence of context-specific coding by combinations of calls, a phenomenon that has been reported from putty-nosed monkeys (*Cercopithecus nictitans martini*) (Arnold and Zuberbühler, 2006a, 2006b) and Campbell's monkeys (Ouattara et al., 2009b).

How do these findings relate to human language? Humans possess an extensive repertoire of speech sounds that can be combined in different ways to convey meanings by adhering to language-specific syntactic, semantic and grammatical rules. Humans use communication as part of their social interactions and dialog mainly consists of exchanging social information about personal experiences and relationships (Dunbar et al., 1997). In contrast, non-human primates have a much more finite repertoire, but nonetheless some species have the ability to combine calls to convey meanings. As an example, colobus monkeys utter specific call combinations depending on the predator type (Schel et al., 2009) and putty-nosed monkeys combine two call types to vary meanings (Arnold and Zuberbühler, 2006b).

Until now, the songs of white-handed gibbons have not been very important in elucidating the transition from primate to human communication, despite the fact that their song system resembles the human speech system more than most other primate call system. Individual notes show pitch modulation and can be assembled into long and sophisticated songs. Within these songs, notes are part of smaller structures, such as 'figures' and 'phrases'. Previous work has already shown that different songs are produced in different contexts, particularly predatory and social events (Clarke et al., 2006). This thesis followed up on this basic finding by providing evidence of a functionally referential alarm calling system in white-handed gibbons. Gibbons have an ability to combine calls in an even more elaborated way than previously thought, increasing the complexity of information that could be conveyed and be beneficial to conspecifics. These remarkable abilities make gibbons, at the basal position of the hominoid family, a relevant model to investigate the origin of human language and whether human and non-human primates coding systems might have shared common syntactic, compositional and/or semantic rules through their evolutionary distant predecessors.

5.4 Future research

This thesis aimed to contribute to a better understand of the singing behaviour of white-handed gibbons and provided new insights into their coding and inference abilities as well as the factors driving non-agressive and competitive interactions with neighbours. Further studies are likely to provide new insights into the problem of how animals avoid or generate signal overlap, with is directly related to the questions about evolution of cooperation and language.

In the first study, I found that duet counter-singing is deployed strategically and most likely as part of territorial behaviour. Future research could address this in more detail, by taking into account distance in a more detailed way, which would require multiple observers. This would allow to quantify that causes and effects of singing on real-time distance fluctuations during turn-taking and counter-singing exchanges and between related and unrelated groups. For example, it would be interesting to see if there is a natural equilibrium point between effective distance and vocal timing respect, i.e. that territorial threat might override the kinship effect, which would predict a change from non-aggressive turn-taking to competitive counter-singing duet song exchanges between kin. Such behaviour was observed during playback experiments, in which we broadcasted natural recorded neighbours duets at close distance. Even though only one of the six tested dyads (target group and stimulus groups) were related, the target group replied by overlapping the 'simulated' duet songs. Furthermore, I also witnessed many times related groups overlapping each other's when tightly spaced (<150m).

Concerning referential signalling, I found structural differences mainly in the early song bout segments, suggesting that gibbons convey meaning in the early stages of their vocal production. Differences in the preliminary song segments differed not only between duet and predator songs, but also between python and leopard songs. My results suggest that the early song segments thus provide informational cues about the type of predator. However, the difference found in the following song segments was largely based on production differences in 'sharp wow' notes, which were given during the louder song sequences of predator songs, suggesting that callers were targeting neighbouring groups.

The later song segments have not been analysed systematically, although some anecdotal observations of waveform and spectrogram features of duet and predator songs suggested that there are further differences. Specifically, the two song types could be grouped based on audibility: the preliminary song segments of low sound level (i.e. soft ‘hoo’ notes series and the following ‘oo’ segments), and the remaining song bout, that gradually increases in volume as the first non ‘hoo’ and ‘oo’ notes start to be produced. Importantly, the two song parts might target different audiences, groupmates and neighbouring individuals. Also, leopard songs were significantly longer than python songs, but this component is unlikely to function as a carrier of meaning. Hence, there is a need to analyse the later parts of gibbon songs to understand what, if anything, is coded in the rest of the song, that could benefit either group members or neighbouring groups.

Finally, further playback experiments will be needed to explore the degree to which the different structures discovered in gibbon song are perceived and understood by recipients, such as broadcasting different predator song types (i.e. leopard, python songs or disturbed songs produced to human presence), or by combining artificially different contextual song structures in incongruent ways (e.g. *Hoo-oo figures* or delayed great call production in duet songs or *Ooaa bursts* in predator songs).

This research provided new insights on the singing behaviour of white-handed gibbon (*Hylobates lar*), and in the process, on the questions in relation to the biological roots of language-related capacities, such as vocal turn-taking, vocal coding and inference abilities, making Hylobatids a noteworthy model to investigate the origin of human language.

6- BIBLIOGRAPHY

Aguilar de Soto, N., Madsen, P.T., Tyack, P., Arranz, P., Marrero, J., Fais, A., Revelli, E., and Johnson, M. (2012). No shallow talk: Cryptic strategy in the vocal communication of Blainville's beaked whales. Mar. Mammal Sci. *28*, E75–E92.

Altmann, S.A., and Altmann, J. (1977). On the analysis of rates of behaviour. Anim. Behav. *25*, 364–372.

Anandam, M.V., Groves, C.P., and Molur, S. (2013). Species accounts of Hylobatidae. In Handbook of the Mammals of the World Primates, (Barcelona, Spain), pp. 778–791.

Andrieu, J., Penny, S.G., Bouchet, H., Malaivijitnond, S., Reichard, U.H., and Zuberbühler, K. (2020). White-handed gibbons discriminate context-specific song compositions. PeerJ, 8, e9477.

Andrieu, J., Neumann, C., Malaivijitnond, S., Reichard, U.H., and Zuberbühler, K. Nepotism enhances intergroup communication in gibbons. in prep a.

Andrieu, J., Penny, S.G., Bouchet, H., Malaivijitnond, S., Reichard, U.H., and Zuberbühler, K. Gibbon songs refer to external events. in prep b.

Arlet, M., Jubin, R., Masataka, N., and Lemasson, A. (2015). Grooming-at-a-distance by exchanging calls in non-human primates. Biol. Lett. *11*, 20150711.

Armstrong, E.A. (1973). A study of bird song. Dover Publ. *201*.

Arnold, K., and Zuberbühler, K. (2006a). Language evolution: Semantic combinations in primate calls. Nature *441*, 303–303.

Arnold, K., and Zuberbühler, K. (2006b). The alarm-calling system of adult male putty-nosed monkeys, *Cercopithecus nictitans martini*. Anim. Behav. *72*, 643–653.

Arnold, K., and Zuberbühler, K. (2013). Female Putty-Nosed Monkeys Use Experimentally Altered Contextual Information to Disambiguate the Cause of Male Alarm Calls. PLoS ONE *8*.

Asensio, N., Brockelman, W.Y., Malaivijitnond, S., and Reichard, U.H. (2014). White-handed Gibbon (*Hylobates lar*) core area use over a short-time scale. Biotropica *46*, 461–469.

Bailey, W.J. (2003). Insect duets: underlying mechanisms and their evolution. Physiol. Entomol. *28*, 157–174.

Bailey, W.J., and Hammond, T.J. (2003). Duetting in insects does call length influence reply latency? J. Zool. *260*, 267–274.

Barelli, C., Heistermann, M., Boesch, C., and Reichard, U.H. (2007). Sexual swellings in wild white-handed gibbon females (*Hylobates lar*) indicate the probability of ovulation. Horm. Behav. *51*, 221–230.

Barelli, C., Boesch, C., Heistermann, M., and Reichard, U.H. (2008a). Female White-handed gibbons (*Hylobates lar*) lead group movements and have priority access to food resources. Behaviour *145*.

Barelli, C., Heistermann, M., Boesch, C., and Reichard, U.H. (2008b). Mating patterns and sexual swellings in pair-living and multimale groups of wild white-handed gibbons, *Hylobates lar*. Anim. Behav. *75*, 991–1001.

Barelli, C., and Heistermann, M. (2012). Sociodemographic correlates of faecal androgen levels in wild male white-handed gibbons (*Hylobates lar*). Int. J. Primatol. *33*, 784–798.

Barelli, C., Mundry, R., Heistermann, M., and Hammerschmidt, K. (2013a). Cues to androgens and quality in male gibbon Songs. PLoS ONE *8*, e82748.

Barelli, C., Matsudaira, K., Wolf, T., Roos, C., Heistermann, M., Hodges, K., Ishida, T., Malaivijitnond, S., and Reichard, U.H. (2013b). Extra-pair paternity confirmed in wild white-handed gibbons. Am. J. Primatol. n/a-n/a.

Bates, Mary.E., and Cropp, Beth.F. (2010). Spatial location influences vocal interactions in bullfrog choruses. J. Acoust. Soc. Am. *127*.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting Linear Mixed-Effects Models using lme4. ArXiv14065823 Stat.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. Ser. B Methodol. *57*, 289–300.

Bergmüller, R., Johnstone, R.A., Russell, A.F., and Bshary, R. (2007). Integrating cooperative breeding into theoretical concepts of cooperation. Behav. Processes *76*, 61–72.

Berthet, M., Mesbahi, G., Pajot, A., Cäsar, C., Neumann, C., and Zuberbühler, K. (2019). Titi monkeys combine alarm calls to create probabilistic meaning. Sci. Adv. *5*, eaav3991.

Boer, L.E.M. de (1972). Marked chromosomes associations in catarrhine monkeys, with a note on chromosome associations in other primate groups. Journal of Human Evolution*, 1*(1), 83-86.

Boissy, A. (1995). Fear and fearfulness in animals. Q. Rev. Biol. *70*, 165–191.

Bolhuis, J.J., Beckers, G.J., Huybregts, M.A., Berwick, R.C., and Everaert, M.B. (2018). Meaningful syntactic structure in songbird vocalizations? PLoS Biol. *16*, e2005157.

Borjon, J.I., and Ghazanfar, A.A. (2014). Convergent evolution of vocal cooperation without convergent evolution of brain size. Brain. Behav. Evol. *84*, 93–102.

Brockelman, W.Y. (1975). Gibbon population and their conservation in Thailand. Nat Hist Bull Siam Soc *26*, 133–157.

Brockelman, W.Y., and Srikosamatara, S. (1984). Maintenance and evolution of social structure in gibbons. In The Lesser Apes: Evolutionary and Behavioural Biology, (Edinburgh: Edinburgh University Press), pp. 298–323.

Brockelman, W.Y., and Srikosamatara, S. (1993). Estimation of density of gibbon groups by use of loud songs. Am. J. Primatol. *29*, 93–108.

Brockelman, W.Y., Reichard, U.H., Treesucon, U., and Raemaekers, J.J. (1998). Dispersal, pair formation and social structure in Gibbons (*Hylobates lar*). Behav. Ecol. Sociobiol. *42*, 329–339.

Brunet, Y., Lagouarde, J.P., and Zouboff, V. (1996). Estimating long-term microclimatic conditions for long-range sound propagation studies. pp. 123–136. In 7. International Symposium on Long Range Sound Propagation.

Calenge, C. (2011). Home range estimation in R: the adehabitat HR package (Saint Benoist, France).

Carbone, L., Alan Harris, R., Gnerre, S., Veeramah, K.R., Lorente-Galdos, B., Huddleston, J., Meyer, T.J., Herrero, J., Roos, C., Aken, B., et al. (2014). Gibbon genome and the fast karyotype evolution of small apes. Nature *513*, 195–201.

Carnap, R. (1988). Meaning and necessity: a study in semantics and modal logic (Vol. 30). University of Chicago Press.

Caro, T. (2005). Antipredator defenses in birds and mammals. University of Chicago Press.

Carpenter, C.R. (1940). A field study in Siam of the behavior and social relations of the gibbon (*Hylobates lar*). Johns Hopkins Press *16*, 80–84.

Cäsar, C., and Zuberbühler, K. (2012). Referential alarm calling behaviour in New World primates. Curr. Zool. *58*, 680–697.

Cäsar, C., Zuberbühler, K., Young, R.J., and Byrne, R.W. (2013). Titi monkey call sequences vary with predator location and type. Biol. Lett. *9*, 20130535.

Catchpole, C.K., and Slater, P.J.B. (2003). Bird song: biological themes and variations (Cambridge University Press).

Chan, B.P.L., Mak, C. fung, Yang, J., and Huang, X. (2017). Population, distribution, vocalization and conservation of the Gaoligong Hoolock Gibbon (*Hoolock tianxing*) in the Tengchong Section of the Gaoligongshan National Nature Reserve, China. Primate Conserv. *31*, 107–113.

Chan, Y.C., Roos, C., Inoue-Murayama, M., Inoue, E., Shih, C.C., Pei, K.J.C., and Vigilant, L. (2010). Mitochondrial genome sequences effectively reveal the phylogeny of Hylobates Gibbons. PLoS ONE *5*, e14419.

Cheyne, S.M., Chivers, D.J., and Sugardjito, J. (2007). Covariation in the great calls of rehabilitant and wild gibbons (*Hylobates albibarbis*). Raffles Bull. Zool. *55*, 201–207.

Chiarelli, B. (1975). The study of primate chromosomes. Primate Funct. Morphol. Evol. 103–127.

Chivers, D.J. (1974). Siamang in Malaya.

Chivers, D.J. (1977). The lesser apes. Primate Conserv. 539–598.

Chivers, D.J., and Gittins, S.P. (1978). Diagnostic features of gibbon species. Int. Zoo Yearb. *18*.

Chowdhury, S.A., Danieli, M., and Riccardi, G. (2015). The role of speakers and context in classifying competition in overlapping speech. In Sixteenth Annual Conference of the International Speech Communication Association, p.

Clarke, E., Reichard, U.H., and Zuberbühler, K. (2006). The syntax and meaning of wild gibbon Songs. PLoS ONE *1*, e73.

Clarke, E., Reichard, U.H., and Zuberbühler, K. (2012). The anti-predator behaviour of wild White-handed gibbons (*Hylobates lar*). Behav. Ecol. Sociobiol. *66*, 85–96.

Clarke, E., Reichard, U.H., and Zuberbühler, K. (2015). Context-specific close-range "hoo" calls in wild gibbons (*Hylobates lar)*. BMC Evol. Biol. *15*, 56.

Clink, D.J., Bernard, H., Crofoot, M.C., and Marshall, A.J. (2017). Investigating individual vocal signatures and small-scale patterns of geographic variation in female Bornean gibbon (*Hylobates muelleri*) great calls. Int. J. Primatol. *38*, 656–671.

Clutton-Brock, T.H., Brotherton, P.N.M., O'Riain, M.J., Griffin, A.S., Gaynor, D., Kansky, R., Sharpe, L., and McIlrath, G.M. (2001). Contributions to cooperative rearing in meerkats. Anim. Behav. *61*, 705–710.

Cowlishaw, G. (1992). Song function in gibbons. Behaviour *121*, 131–153.

Dabelsteen, T., McGregor, P.K., Shepherd, M., Whittaker, X., and Pedersen, S.B. (1996). Is the signal value of overlapping different from that of alternating during matched singing in great tits? J. Avian Biol. *27*, 189–194.

Dabelsteen, T., McGregor, P.K., Holland, J., Tobias, J.A., and Boel Pedersen, S. (1997). The signal function of overlapping singing in male robins. Anim. Behav. *53*, 249–256.

Deng, H., Zhou, J., and Yang, Y. (2014). Sound spectrum characteristics of songs of Hainan gibbon (*Nomascus hainanu*s). Int. J. Primatol. *35*, 547–556.

Deng, H., Gao, K., and Zhou, J. (2016). Non-specific alarm calls trigger mobbing behavior in Hainan gibbons (*Nomascus hainanus*). Sci. Rep. *6*, 34471.

Dooley, H.M., Judge, D.S., and Schmitt, L.H. (2013). Singing by male and female Kloss gibbons (*Hylobates klossii*) in the Peleonan Forest, Siberut Island, Indonesia. Primates *54*, 39–48.

Dunbar, R.I.M., Marriott, A., and Duncan, N.D.C. (1997). Human conversational behavior. Hum. Nat. *8*, 231–246.

Dunn, O.J. (1964). Multiple comparisons using rank sums. Technometrics *6*, 241–252.

Ellefson, J.O. (1968). Territorial behavior in the common white-handed gibbon (*Hylobates lar*). pp. 180–199.

Engqvist, L. (2005). The mistreatment of covariate interaction terms in linear model analyses of behavioural and evolutionary ecology studies. Anim. Behav. *70*.

Fairbanks, L.A. (1990). Reciprocal benefits of allomothering for female vervet monkeys. Anim. Behav. *40*, 553–562.

Fan, P., Liu, C., Luo, W., and Jiang, X. (2006). Can a group elicit duets from its neighbours? A field study on the Black-Crested Gibbon (*Nomascus concolor jingdongensis*) in Central Yunnan, China. Folia Primatol. (Basel) *78*, 186–195.

Fan, P.-F., Ai, H.-S., Fei, H.-L., Zhang, D., and Yuan, S.-D. (2013). Seasonal variation of diet and time budget of Eastern hoolock gibbons (*Hoolock leuconedys*) living in a northern montane forest. Primates *54*, 137–146.

Fan, P.-F., He, K., Chen, X., Ortiz, A., Zhang, B., Zhao, C., Li, Y.-Q., Zhang, H.-B., Kimock, C., Wang, W.-Z., et al. (2017). Description of a new species of Hoolock gibbon (Primates: Hylobatidae) based on integrative taxonomy. Am. J. Primatol. *79*, e22631.

Farabaugh, S.M. (1982). The ecological and social significance of duetting. In Acoustic Communication in Birds, (DE Kroodsma and EH Miller), pp. 85–124.

Farabaugh, S.M. (1985). A comparative study of duet song in tropical Thryothorus wrens.

Fedurek, P., Zuberbühler, K., and Dahl, C.D. (2016). Sequential information in a great ape utterance. Sci. Rep. *6*, 38226.

Feng, J.-J., Cui, L.-W., Ma, C.-Y., Fei, H.-L., and Fan, P.-F. (2014). Individuality and stability in male songs of cao vit gibbons (*Nomascus nasutus*) with potential to monitor population dynamics. PLoS One *9*, e96317.

Fichtel, C. (2008). Ontogeny of conspecific and heterospecific alarm call recognition in wild Verreaux’s sifakas (*Propithecus verreauxi verreauxi*). Am. J. Primatol. Off. J. Am. Soc. Primatol. *70*, 127–135.

Ficken, R.W., Ficken, M.S., and Hailman, J.P. (1974). Temporal pattern shifts to avoid acoustic interference in singing birds. Science *183*, 762–763.

Fischer, J., Hammerschmidt, K., and Todt, D. (1995). Factors affecting acoustic variation in Barbary-macaque (*Macaca sylvanus*) disturbance calls. Ethology *101*, 51–66.

Fischer, J., Hammerschmidt, K., and Todt, D. (1998). Local variation in Barbary macaque shrill barks. Anim. Behav. *56*, 623–629.

Fischer, J., Hammerschmidt, K., Cheney, D.L., and Seyfarth, R.M. (2001a). Acoustic features of female Chacma baboon barks. Ethology *107*, 33–54.

Fischer, J., Metz, M., Cheney, D.L., and Seyfarth, R.M. (2001b). Baboon responses to graded bark variants. Anim. Behav. *61*, 925–931.

Fischer, J., Kitchen, D.M., Seyfarth, R.M., and Cheney, D.L. (2004). Baboon loud calls advertise male quality: acoustic features and their relation to rank, age, and exhaustion. Behav. Ecol. Sociobiol. *56*, 140–148.

Fischer, J., and Hammerschmidt, K. (2006). Vocal communication in Barbary macaques: a comparative perspective. Barbary Macaque Biol. Manag. Conserv. 63.

Fitch, W.T. (2004). Kin selection and 'mother tongues': a neglected component in language evolution. Evol. Commun. Syst. Comp. Approach 275–296.

Fitch, W.T. (2005). The Evolution of Language: A Comparative Review. Biol. Philos. *20*, 193–203.

Fitzsimmons, L.P., Foote, J.R., Ratcliffe, L.M., and Mennill, D.J. (2008). Frequency matching, overlapping and movement behaviour in diurnal countersinging interactions of black-capped chickadees. Anim. Behav. *75*, 1913–1920.

Geissmann, T. (1993). Evolution of communication in gibbons (Hylobatidae) (Ph. D. thesis, Anthropological Institute, Philosoph. Faculty II, Zürich University. 374 pp.(English text, German summary)).

Geissmann, T. (1995). Gibbon systematics and species identification. Int. Zoo News *42*, 467–501.

Geissmann, T. (1999a). Duet songs of the siamang, *Hylobates syndactylus*: II. Testing the pair-bonding hypothesis during a partner exchange. Behaviour *136*, 1005–1039.

Geissmann, T. (1999b). Duet songs of the siamang, *Hylobates syndactylus*: II. Testing the pair-bonding hypothesis during a partner exchange. Behaviour *136*, 1005–1039.

Geissmann, T. (2000a). Gibbon song and human music from an evolutionary perspective. In The Origins of Music, (Cambridge, Massachusetts: MIT Press), pp. 103–123.

Geissmann, T. (2000b). Duet songs of the Siamang *Hylobates syndactylus*: I. Structure and organisation. Primate Rep. *56*, 33–60.

Geissmann, T. (2002a). Duet-splitting and the evolution of gibbon songs. Biol Rev *77*, 57–76.

Geissmann, T. (2002b). Taxonomy and evolution of gibbons. Evol. Anthropol. Issues News Rev. 28–31.

Geissmann, T. (2007). Status reassessment of the gibbons: results of the Asian primate red list workshop 2006. Gibbon J. 5–15.

Geissmann, T., and Nijman, V. (2006). Calling in wild silvery gibbons (*Hylobates moloch*) in Java (Indonesia): behavior, phylogeny, and conservation. Am. J. Primatol. *68*, 1–19.

Geissmann, T., and Orgeldinger, M. (1997). Pair bond and duet songs in siamangs (*Hylobates syndactylus*)—. Adv Ethol *32*, 123.

Geissmann, T., and Orgeldinger, M. (2000). The relationship between duet songs and pair bonds in siamangs, *Hylobates syndactylus*. Anim. Behav. *60*, 805–809.

Geissmann, T., Bohlen-Eyring, S., and Heuck, A. (2005). The male song of the Javan silvery gibbon (*Hylobates moloch*). Contrib. Zool. *74*, 1–25.

Gerhardt, H.C., and Huber, F. (2002). Acoustic communication in Insects and Anurans: common problems and diverse solutions (University of Chicago Press).

Ghazanfar, A.A., and Takahashi, D.Y. (2014). The evolution of speech: vision, rhythm, cooperation. Trends Cogn. Sci. *18*, 543–553.

Gil, D., and Gahr, M. (2002). The honesty of bird song: multiple constraints for multiple traits. ScienceDirect. Trends Ecol. Evol. *17*.

Gittins, S.P., and Raemaekers, J.J. (1980). Siamang, Lar and Agile Gibbons. (University of Cambridge), pp. 63–105.

Grassman, L.I. (2001). Spatial ecology and conservation of the felid community in Phu Khieo Wildlife Sanctuary, Thailand. Rep. Cat Action Treas.

Greenfield, M.D. (1994a). Cooperation and conflict in the evolution of signal interactions. Annu. Rev. Ecol. Syst. *25*, 97–126.

Greenfield, M.D. (1994b). Synchronous and alternating choruses in Insects and Anurans: common mechanisms and diverse functions. Am. Zool. *34*, 605–615.

Grice, H.P. (1969). Utterer’s meaning and intentions. Philos. Rev. *78*, 147–177.

Groves, C. (2001). Primate Taxonomy (Smithsonian Books).

Grow, N.B. (2019). Cryptic communication in a montane nocturnal Haplorhine, *Tarsius pumilus*. Folia Primatol. (Basel) *90*, 404–421.

Haimoff, E.H. (1981). Video analysis of Siamang (Hylobates Syndactylus) songs. Behaviour *76*, 128–151.

Haimoff, E. (1983). Gibbon songs: An acoustical, organizational and behavioural analysis. Ph.D. University of Cambridge.

Haimoff, E.H. (1984a). Acoustic and organizational features of gibbon songs. In The Lesser Apes: Evolutionary and Behavioural Biology, (Edinburgh: H. Preuschoft, D. J. Chivers, W. Y. Brockelman and N. Creel), pp. 333–353.

Haimoff, E.H. (1984b). The organization of song in the Agile Gibbon (*Hylobates agilis*). Folia Primatol. (Basel) *42*, 42–61.

Haimoff, E.H. (1984c). The organization of song in the Hainan black gibbon (*Hylobates concolor hainanus*). Primates *25*, 225–235.

Haimoff, E.H. (1986). Convergence in the duetting of monogamous Old World primates. J. Hum. Evol. *15*, 51–59.

Haimoff, E.H., and Gittins, S.P. (1985). Individuality in the songs of wild agile gibbons (*Hylobates agilis*) of Peninsular Malaysia. Am. J. Primatol. *8*, 239–247.

Hall, M.L. (2004). A review of hypotheses for the functions of avian duetting. Behav. Ecol. Sociobiol. *55*, 415–430.

Hall, M.L. (2009). A review of vocal duetting in Birds. In Advances in the Study of Behavior, (Elsevier), pp. 67–121.

Hall, M.L., Illes, A., and Vehrencamp, S.L. (2006). Overlapping signals in banded wrens: long-term effects of prior experience on males and females. Behav. Ecol. *17*, 260–269.

Haraway, M.M., and Maples, E.G. (1998). Flexibility in the species-typical songs of gibbons. Primates *39*, 1–12.

Harrison, T. (2016). The fossil record and evolutionary history of Hylobatids. In Evolution of Gibbons and Siamang: Phylogeny, Morphology, and Cognition, U.H. Reichard, H. Hirai, and C. Barelli, eds. (New York, NY: Springer), pp. 91–110.

Helfer, B., and Osiejuk, T.S. (2015). It takes all kinds in acoustic communication: a new perspective on the song overlapping phenomenon. Ethology *121*, 315–326.

Inoue, Y., Sinun, W., Yosida, S., and Okanoya, K. (2013). Intergroup and intragroup antiphonal songs in wild male Mueller's gibbons (*Hylobates muelleri*). Interact. Stud. *14*, 24–43.

Inoue, Y., Sinun, W., and Okanoya, K. (2016). Activity budget, travel distance, sleeping time, height of activity and travel order of wild East Bornean Grey gibbons (*Hylobates funereus*) in Danum Valley Conservation Area. Raffles Bull. Zool. *64*.

Kappeler, M. (1981). The Javan Silvery Gibbon (Hylobates lar moloch). Universitat Basel.

Keenan, S., Lemasson, A., and Zuberbühler, K. (2013). Graded or discrete? A quantitative analysis of Campbell's monkey alarm calls. Anim. Behav. *85*, 109–118.

Kenyon, M., Roos, C., Binh, V.T., and Chivers, D. (2011). Extrapair paternity in Golden-Cheeked Gibbons (*Nomascus gabriellae*); in the secondary lowland forest of Cat Tien National Park, Vietnam. Folia Primatol. (Basel) *82*, 154–164.

Kershenbaum, A. (2014). Entropy rate as a measure of animal vocal complexity. Bioacoustics *23*, 195–208.

Kershenbaum, A., and Garland, E.C. (2015). Quantifying similarity in animal vocal sequences: Which metric performs best? Methods Ecol. Evol. *6*, 1452–1461.

Kershenbaum, A., Bowles, A.E., Freeberg, T.M., Jin, D.Z., Lameira, A.R., and Bohn, K. (2014). Animal vocal sequences: not the Markov chains we thought they were. Proc. R. Soc. B Biol. Sci. *281*, 20141370.

Kershenbaum, A., Blumstein, D.T., Roch, M.A., Akçay, Ç., Backus, G., Bee, M.A., Bohn, K., Cao, Y., Carter, G., and Cäsar, C. (2016). Acoustic sequences in non-human animals: a tutorial review and prospectus. Biol. Rev. *91*, 13–52.

Kirchhof, J., and Hammerschmidt, K. (2006). Functionally referential alarm calls in tamarins (*Saguinus fuscicollis* and *Saguinus mystax*) – Evidence from playback experiments. Ethology *112*, 346–354.

Kitchen, D.M., Cheney, D.L., Seyfarth, R.M., and Beehner, J.C. (2009). A link between faecal testosterone and an honest signal-the loud 'wahoo' vocalizations of chacma baboons. Am. J. Phys. Anthropol. 166–166.

Koda, H., Lemasson, A., Oyakawa, C., Pamungkas, J., and Masataka, N. (2013). Possible role of mother-daughter vocal interactions on the development of species-specific song in gibbons. PloS One *8*, e71432.

Koenig, A. (1998). Visual scanning by common marmosets (*Callithrix jacchus*): functional aspects and the special role of adult males. Primates *39*, 85–90.

Konrad, R., and Geissmann, T. (2006). Vocal diversity and taxonomy of Nomascus in Cambodia. Int. J. Primatol. *27*, 713–745.

Kumar, A., Devi, A., Gupta, A. K., and Sarma, K. (2013). Population, behavioural ecology and conservation of Hoolock Gibbon in Northeast India. In Rare Animals of India, (Singaravelan Natarajan BENTHAM SCIENCE PUBLISHERS), pp. 242–266.

Kunc, H.P., Amrhein, V., and Naguib, M. (2006). Vocal interactions in nightingales, *Luscinia megarhynchos*: more aggressive males have higher pairing success. Anim. Behav. *72*, 25–30.

Lemasson, A., Ouattara, K., Bouchet, H., and Zuberbühler, K. (2010). Speed of call delivery is related to context and caller identity in Campbell's monkey males. Naturwissenschaften *97*, 1023–1027.

Lemasson, A., Glas, L., Barbu, S., Lacroix, A., Guilloux, M., Remeuf, K., and Koda, H. (2011). Youngsters do not pay attention to conversational rules: is this so for non-human primates? Sci. Rep. *1*.

Lenth, R.V. (2016). Least-Squares Means: The R Package lsmeans. J. Stat. Softw. *069*.

Levinson, S.C. (2016). Turn-taking in Human communication – Origins and implications for language processing. Trends Cogn. Sci. *20*, 6–14.

Lukas, D., and Clutton-Brock, T. (2018). Social complexity and kinship in animal societies. Ecol. Lett. *21*, 1129–1134.

MacKinnon, J., and MacKinnon, K. (1980). The behavior of wild spectral tarsiers. Int. J. Primatol. *1*, 361–379.

Manser, M.B. (2009). What do functionally referential alarm calls refer to? In Cognitive Ecology II, pp. 229–246.

Marler, P., Tenaza, R., and Sebeok, T.A. (1977). How animals communicate (Indiana University Press, Bloomington).

Marshall, J.T., and Marshall, E.R. (1976). Gibbons and their territorial songs. Science *193*, 235–237.

Marshall, J.T., and Sugardjito, J. (1986). Gibbon systematics. In Comparative Primate Biology, Vol. 1: Systematics, Evolution and Anatomy, (New York: D. R. Swindler and J. Erwin), pp. 137–185.

Matsudaira, K., and Ishida, T. (2010). Phylogenetic relationships and divergence dates of the whole mitochondrial genome sequences among three gibbon genera. Mol. Phylogenet. Evol. *55*, 454–459.

Matsudaira, K., Ishida, T., Malaivijitnond, S., and Reichard, U.H. (2018). Short dispersal distance of males in a wild white-handed gibbon (*Hylobates lar*) population. Am. J. Phys. Anthropol. *1*.

Matsui, A., Rakotondraparany, F., Munechika, I., Hasegawa, M., and Horai, S. (2009). Molecular phylogeny and evolution of prosimians based on complete sequences of mitochondrial DNAs. Gene *441*, 53–66.

McGregor, P.K. (2005). Animal ommunication Networks (Cambridge University Press).

McGregor, P.K., Otter, and Peake, T.M. (2000). Communication networks: receiver and signaller perspectives. Animal signals: signalling and signal design in animal communication.

Méndez Cardenas, M.G., and Zimmermann, E. (2009). Duetting-A mechanism to strengthen pair bonds in a dispersed pair-living primate (*Lepilemur edwardsi*)? Am. J. Phys. Anthropol. *139*, 523–532.

Mennill, D.J., and Ratcliffe, L.M. (2004a). Do male black-capped chickadees eavesdrop on song contests? A multi-speaker playback experiment. Behaviour *141*, 125–139.

Mennill, D.J., and Ratcliffe, L.M. (2004b). Overlapping and matching in the song contests of black-capped chickadees. Anim. Behav. *67*, 441–450.

Mitani, J.C. (1985a). Responses of gibbons (*Hylobates muelleri*) to self, neighbor, and stranger song duets. Int. J. Primatol. *6*, 193–200.

Mitani, J.C. (1985b). Gibbon song duets and intergroup spacing. Behaviour 59–96.

Mitani, J.C. (1985c). Location-specific responses of gibbons (*Hylobates muelleri*) to male songs. Z. Für Tierpsychol. *70*, 219–224.

Mittermeier, R.A., Rylands, A.B., and Wilson, D.E. (2013). Handbook of the Mammals of the World Primates (Barcelona, Spain).

Mootnick, A., and Fan, P.-F. (2011). A comparative study of Crested gibbons (Nomascus). Am. J. Primatol. 135–154.

Mootnick, A., and Groves, C. (2005). A New generic name for the Hoolock gibbon (Hylobatidae). Int. J. Primatol. *26*, 971–976.

Morino, L. (2010). Clouded leopard predation on a wild juvenile Siamang. Folia Primatol. (Basel) *81*, 362–368.

Müller, A.E., and Anzenberger, G. (2002). Duetting in the titi Monkey *Callicebus cupreus*: structure, pair specificity and development of duets. Folia Primatol. (Basel) *73*, 104–115.

Mundry, R., and Fischer, J. (1998). Use of statistical programs for nonparametric tests of small samples often leads to incorrect Pvalues: examples from Animal Behaviour. Anim. Behav. *56*, 256–259.

Naguib, M. (1999). Effects of song overlapping and alternating on nocturnally singing nightingales. Anim. Behav. *58*, 1061–1067.

Naguib, M., and Todt, D. (1997). Effects of dyadic vocal interactions on other conspecific receivers in nightingales. Anim. Behav. *54*, 1535–1544.

Naguib, M., and Wiley, R.H. (2001). Estimating the distance to a source of sound: mechanisms and adaptations for long-range communication. Anim. Behav. *62*, 825–837.

Naguib, M., and Mennill, D.J. (2010). The signal value of birdsong: empirical evidence suggests song overlapping is a signal. Anim. Behav. *80*, e11–e15.

Narins, P.M. (1992). Evolution of Anuran chorus behavior: neural and behavioral constraints. Am. Nat. *139*, S90–S104.

Neudenberger, J. (1993). Monogamie als Paarungssystem: Eine Fallstudie am Weißhandgibbon (Hylobates lar) im Khao Yai Nationalpark, Thailand. Unveröff. Diplomarb.

Ngoprasert, D., and Gale, G.A. (2019). Tiger density, dhole occupancy, and prey occupancy in the human disturbed Dong Phayayen – Khao Yai Forest Complex, Thailand. Mamm. Biol. *95*, 51–58.

Nietsch, A. (1999). Duet vocalizations among different populations of Sulawesi tarsiers. Int. J. Primatol. *20*, 567–583.

Nowicki, S., and Searcy, W.A. (2014). The evolution of vocal learning. Curr. Opin. Neurobiol. *28*, 48–53.

Orgeldinger, M. (1997). Protective and territorial behavior in captive siamangs (*Hylobates syndactylus*). Zoo Biol. *16*, 309–325.

Ouattara, K., Zuberbühler, K., N'goran, E.K., Gombert, J.-E., and Lemasson, A. (2009a). The alarm call system of female Campbell's monkeys. Anim. Behav. *78*, 35–44.

Ouattara, K., Lemasson, A., and Zuberbühler, K. (2009b). Campbell's monkeys concatenate vocalizations into context-specific call sequences. Proc. Natl. Acad. Sci. *106*, 22026–22031.

Ouattara, K., Lemasson, A., and Zuberbühler, K. (2009c). Campbell's monkeys Use affixation to alter call meaning. PLoS ONE *4*, e7808.

Oyakawa, C., Koda, H., and Sugiura, H. (2007). Acoustic features contributing to the individuality of wild Agile gibbon (*Hylobates agilis agilis*) songs. Am. J. Primatol. *69*, 777–790.

Pika, S., Wilkinson, R., Kendrick, K.H., and Vernes, S.C. (2018). Taking turns: bridging the gap between human and animal communication. Proc. Biol. Sci. *285*.

Pozzi, L., Hodgson, J.A., Burrell, A.S., Sterner, K.N., Raaum, R.L., and Disotell, T.R. (2014). Primate phylogenetic relationships and divergence dates inferred from complete mitochondrial genomes. Mol. Phylogenet. Evol. *75*, 165–183.

Prouty, L.A., Buchanan, P.D., Pollitzer, W.S., and Mootnick, A.R. (1983). Taxonomic note: Bunopithecus: A genus-level taxon for the hoolock gibbon (Hylobates hoolock). Am. J. Primatol. *5*, 83–87.

R core Team (2018). R: A language and environment for statistical computing. (Vienna, Austria: R Foundation for Statistical Computing).

Rabinowitz, A., Andau, P., and Chai, P.P.K. (1987). The Clouded leopard in Malaysian Borneo. Oryx *21*, 107–111.

Raemaekers, J.J., Raemaekers, P.M., and Haimoff, E.H. (1984a). Loud calls of the gibbon (*Hylobates lar*): repertoire, organisation and context. Behaviour *91*, 146–189.

Raemaekers, J.J., and Raemaekers, P.M. (1984b). The Ooaa duet of the gibbon (*Hylobates lar*). Folia Primatol. (Basel) *42*, 209.215.

Raemaekers, J.J., and Raemaekers, P.M. (1985a). Field playback of loud calls to gibbons (*Hylobates lar*): territorial, sex-specific and species-specific responses. Anim. Behav. *33*, 481–493.

Raemaekers, P.M., and Raemaekers, J.J. (1985b). Long-range vocal interactions between groups of gibbons (*Hylobates lar*). Behaviour *95*, 26–44.

Rasoloharijaona, S., Randrianambinina, B., Braune, P., and Zimmermann, E. (2006). Loud calling, spacing, and cohesiveness in a nocturnal primate, the Milne Edwards' sportive lemur (*Lepilemur edwardsi*). Am. J. Phys. Anthropol. *129*, 591–600.

Ravignani, A., Verga, L., and Greenfield, M.D. (2019). Interactive rhythms across species: the evolutionary biology of animal chorusing and turn-taking. Ann. N. Y. Acad. Sci. *1453*, 12–21.

Reby, D., Joachim, J., Lek, S., and Aulagnier, S. (1998). Individuality in the groans of fallow deer (*Dama dama*) bucks. J. Zool. *245*, 79–84.

Rehberg-Besler, N., Doucet, S.M., and Mennill, D.J. (2017). Overlapping vocalizations produce far-reaching choruses: a test of the signal enhancement hypothesis. Behav. Ecol. *28*, 494–499.

Reichard, U.H. (1995). Extra-pair copulations in a monogamous gibbon (*Hylobates la*r). Ethology *100*, 99–112.

Reichard, U.H., and Sommer, V. (1997). Group encounters in wild gibbons (*Hylobates lar*): Agonism, affiliation, and the concept of infanticide. Behaviour *134*, 1135–1174.

Reichard, U.H. (1998). Sleeping sites, sleeping places, and preseleep behaviour of gibbons (*Hylobates lar*). Am. J. Primatol. *46*, 35–62.

Reichard, U.H., and Barelli, C. (2008). Life History and Reproductive Strategies of Khao Yai *Hylobates lar*: Implications for Social Evolution in Apes. Int. J. Primatol. *29*, 823–844.

Reichard, U.H. (2009). The social organization and mating system of Khao Yai White-handed gibbons: 1992-2006. In The Gibbons: New Perspectives on Small Ape Socioecology and Population Biology, D. Whittaker, and S. Lappan, eds. (New York, NY: Springer New York), pp. 347–384.

Reichard, U.H., Ganpanakngan, M., and Barelli, C. (2012). White-handed gibbons of Khao Yai: social flexibility, complex reproductive strategies, and a slow life history. In Long-Term Field Studies of Primates, P.M. Kappeler, and D.P. Watts, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 237–258.

Reichard, U.H., Hirai, H., and Barelli, C. (2016). Evolution of Gibbons and Siamang: Phylogeny, Morphology, and Cognition (Springer).

Robinson, J.G. (1981). Vocal regulation of inter- and intragroup spacing during boundary encounters in the titi monkey, *Callicebus moloch*. Primates *22*, 161–172.

Roos, C. (2016). Phylogeny and Classification of Gibbons (Hylobatidae). In Evolution of Gibbons and Siamang, (Springer, New York, NY), pp. 151–165.

Rossano, F. (2013). Sequence organization and timing of bonobo mother-infant interactions. Interact. Stud. *14*, 160–189.

Rossano, F., and Liebal, K. (2014). “Requests” and “offers” in orang-utans and human infants. In Requesting in Social Interaction, (Drew, P. and Couper-Kuhlen, E.), pp. 333–362.

Sacks, H., Schegloff, E.A., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. language *50*, 696–735.

Savini, T., Boesch, C., and Reichard, U.H. (2008). Home-range characteristics and the influence of seasonality on female reproduction in white-handed gibbons (*Hylobates lar*) at Khao Yai National Park, Thailand. Am. J. Phys. Anthropol. *135*, 1–12.

Savini, T., Boesch, C., and Reichard, U.H. (2009). Varying ecological quality influences the probability of polyandry in White-handed Gibbons (*Hylobates lar*) in Thailand: polyandry in White-handed Gibbons. Biotropica *41*, 503–513.

Schel, A.M., Tranquilli, S., and Zuberbühler, K. (2009). The alarm call system of two species of black-and-white Colobus monkeys (*Colobus polykomos* and *Colobus guereza*). J. Comp. Psychol. *123*, 136–150.

Schultz, A.H. (1933). Observations on the growth, classification and evolutionary specialization of gibbons and siamangs. Hum. Biol. *5*, 212.

Scott-Phillips, T.C. (2010). Animal communication: insights from linguistic pragmatics. Anim. Behav. *79*, e1–e4.

Searcy, W.A., and Beecher, M.D. (2009a). Song as an aggressive signal in songbirds. Anim. Behav. *78*, 1281–1292.

Searcy, W.A., and Beecher, M.D. (2011). Continued scepticism that song overlapping is a signal. Anim. Behav. *81*, e1–e4.

Seyfarth, R., and Cheney, D. (2018). Pragmatic flexibility in primate vocal production. Curr. Opin. Behav. Sci. *21*, 56–61.

Seyfarth, R.M., Cheney, D.L., and Marler, P. (1980a). Monkey responses to three different alarm calls: evidence of predator classification and semantic communication. Science *210*, 801–803.

Seyfarth, R.M., Cheney, D.L., and Marler, P. (1980b). Vervet monkey alarm calls: semantic communication in a free-ranging Primate. Anim. Behav. *28*, 1070–1094.

Siegel, S., and Castellan, N.J. (1988). The Friedman two-way analysis of variance by ranks. Nonparametric Stat. Behav. Sci. 174–184.

Slocombe, K.E., and Zuberbühler, K. (2005). Functionally referential communication in a chimpanzee. Curr. Biol. *15*, 1779–1784.

Smith, W.J. (1977). The behavior of communicating: an ethological approach Harvard University Press. Camb. MA.

Sommer, V., and Reichard, U. (2000). 14• Rethinking monogamy: the gibbon case. Primate Males Causes Consequences Var. Group Compos. 159.

Stephan, C., and Zuberbühler, K. (2008). Predation increases acoustic complexity in primate alarm calls. Biol. Lett. *4*, 641–644.

Stephan, C., and Zuberbühler, K. (2014). Predation affects alarm call usage in female Diana monkeys (*Cercopithecus diana diana*). Behav. Ecol. Sociobiol. *68*, 321–331.

Stokes, A.W., and Williams, H.W. (1968). Antiphonal calling in Quail. The Auk *85*, 83–89.

Symmes, D., and Biben, M. (1988). Conversational vocal exchanges in squirrel monkeys. In Primate Vocal Communication, P.D.D. Todt, D.P. Goedeking, and D.D. Symmes, eds. (Springer Berlin Heidelberg), pp. 123–132.

Takahashi, D.Y., Narayanan, D.Z., and Ghazanfar, A.A. (2013). Coupled oscillator dynamics of vocal turn-taking in monkeys. Curr. Biol. *23*, 2162–2168.

Tenaza, R.R. (1976). Songs, choruses and countersinging of Kloss' gibbons (*Hylobates klossii*) in Siberut Island, Indonesia. Z. Für Tierpsychol. *40*, 37–52.

Tenaza, R.R., and Tilson, R.L. (1977). Evolution of long-distance alarm calls in Kloss's gibbon. Nature *268*, 233–235.

Terleph, T.A., Malaivijitnond, S., and Reichard, U.H. (2015). Lar gibbon (*Hylobates lar*) great call reveals individual caller identity: Lar Gibbon Great Calls. Am. J. Primatol. *77*, 811–821.

Terleph, T.A., Malaivijitnond, S., and Reichard, U.H. (2016). Age related decline in female lar gibbon great call performance suggests that call features correlate with physical condition. BMC Evol. Biol. *16*, 4.

Terleph, T.A., Malaivijitnond, S., and Reichard, U.H. (2017). Male white-handed gibbons flexibly time duet contributions. Behav. Ecol. Sociobiol. *72*.

Thinh, V.N., Mootnick, A.R., Geissmann, T., Li, M., Ziegler, T., Agil, M., Moisson, P., Nadler, T., Walter, L., and Roos, C. (2010a). Mitochondrial evidence for multiple radiations in the evolutionary history of small apes. BMC Evol. Biol. *10*, 74.

Thinh V. N., Thanh, V.N., and Roos, C. (2010b). A new species of crested gibbon, from the central Annamite mountain range. Vietnam. J. Primatol. *1*, 1–12.

Thinh, V.N., Rawson, B., Hallam, C., Kenyon, M., Nadler, T., Walter, L., and Roos, C. (2010c). Phylogeny and distribution of crested gibbons (genus *Nomascus*) based on mitochondrial cytochrome b gene sequence data. Am. J. Primatol. *72*, 1047–1054.

Thorpe, W.H. (1961). Bird-song: the biology of vocal communication and expression in birds (Oxford, England: University Press).

Thorpe, W.H., Hall-Craggs, J., Hooker, B., Hooker, T., and Hutchinson, R. (1972). Duetting and antiphonal song in Birds: its extent and significance. Behav. Suppl. *III*, 1–197.

Tobias, M.L., Viswanathan, S.S., and Kelley, D.B. (1998). Rapping, a female receptive call, initiates male–female duets in the South African clawed frog. Proc. Natl. Acad. Sci. *95*, 1870–1875.

Todt, D., and Naguib, M. (2000). Vocal interactions in Birds: the use of song as a model in communication. In Advances in the Study of Behavior, P.J.B. Slater, J.S. Rosenblatt, C.T. Snowdon, and T.J. Roper, eds. (Academic Press), pp. 247–296.

Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. Behav. Brain Sci. *28*, 675–691.

Tomasello, M., and Call, J. (2007). The gestural communication of apes and monkeys (New York, NY: Taylor & Francis Group/Lawrence Erlbaum Associates).

Townsend, S.W., Zöttl, M., and Manser, M.B. (2011). All clear? Meerkats attend to contextual information in close calls to coordinate vigilance. Behav. Ecol. Sociobiol. *65*, 1927–1934.

Townsend, S.W., Koski, S.E., Byrne, R.W., Slocombe, K.E., Bickel, B., Boeckle, M., Braga Goncalves, I., Burkart, J.M., Flower, T., and Gaunet, F. (2017). Exorcising Grice's ghost: an empirical approach to studying intentional communication in animals. Biol. Rev. *92*, 1427–1433.

Townsend, S.W., Engesser, S., Stoll, S., Zuberbühler, K., and Bickel, B. (2018). Compositionality in animals and humans. PLoS Biol. *16*, e2006425.

Traeholt, C., Bonthoeun, R., Virak, C., Samuth, M., and Vutthin, S. (2006). Song Activity of the Pileated Gibbon, *Hylobates pileatus*, in Cambodia. Primate Conserv. 139–144.

Uhde, N.L., and Sommer, V. (2002). Antipredatory behavior in gibbons (*Hylobates lar*, Khao Yai, Thailand). In Khao Yai, pp. 168–291.

Uster, D., and Zuberbühler, K. (2001). The functional significance of Diana monkey "clear" calls. Behaviour *138*, 741–756.

Van Gulik, R.H. (1967). The gibbon in China: an essay in Chinese animal lore (Brill).

Van Thien, N., Anh, N.Q.H., Van Ngoc Thinh, L.V.K., and Roos, C. (2017). Distribution of the northern yellow-cheeked gibbon (*Nomascus annamensis*) in central Vietnam. Vietnam. J. Primatol. 83.

Von Helversen, D. (1980). Structure and function of antiphonal duets. Acta XVIII Int Orn Congr.

Wacewicz, S., and Żywiczyński, P. (2018). Language origins: Fitness consequences, platform of trust, cooperation, and turn-taking. Interact. Stud. *19*, 167–182.

Walker, S., Molur, S., Brockelman, W.Y., Das, J., Islam, A., Geissmann, T., and Peng-Fei, F. (2007). Western Hoolock Gibbon. Hoolock Hoolock Harlan 1831.

Wanelik, K.M., Azis, A., and Cheyne, S.M. (2013). Note-, phrase- and song-specific acoustic variables contributing to the individuality of male duet song in the Bornean southern gibbon (*Hylobates albibarbis*). Primates *54*, 159–170.

Weisberg, S. (2001). Yeo-Johnson power transformations. Dep. Appl. Stat. Univ. Minn. Retrieved June *1*, 2003.

Whitington, C., and Treesucon, U. (1991). Selection and treatment of food plants by white-handed gibbons (*Hylobates lar*) in Khao Yai National Park, Thailand. Nat Hist Bull Siam Soc *39*, 111–122.

Wickler, W. (1980). Vocal duetting and the Pair Bond. Z. Für Tierpsychol. *52*, 201–209.

Widdig, A. (2007). Paternal kin discrimination: the evidence and likely mechanisms. Biol. Rev. *82*, 319–334.

Williams, J. (2017). Wind suppresses calling in northern buff-cheeked crested gibbons (*Nomascus annamensis*). The Human Voyage, *1*.

Wilson, E.O. (1975). Sociobiology: the new synthesis (Cambridge).

Wilson, D.R., Ratcliffe, L.M., and Mennill, D.J. (2016). Black-capped chickadees, *Poecile atricapillus*, avoid song overlapping: evidence for the acoustic interference hypothesis. Anim. Behav. *114*, 219–229.

Yin, L.Y., Fei, H.L., Chen, G.S., Li, J.H., Cui, L.W., and Fan, P.F. (2016). Effects of group density, hunting, and temperature on the singing patterns of eastern hoolock gibbons (*Hoolock leuconedys*) in Gaoligongshan, Southwest China. Am. J. Primatol. *78*, 861–871.

Yoshida, S., and Okanoya, K. (2005). Evolution of turn-taking: a bio-cognitive perspective. Cogn. Stud. *12*, 153–165.

Zichello, J.M. (2018). Look in the trees: Hylobatids as evolutionary models for extinct hominins. Evol. Anthropol. Issues News Rev. *27*, 142–146.

Zuberbühler, K., Noë, R., and Seyfarth, R.M. (1997). Diana monkey long-distance calls: messages for conspecifics and predators. Anim. Behav. *53*, 589–604.

Zuberbühler, K., Jenny, D., and Bshary, R. (1999). The predator deterrence function of primate alarm calls. Ethology *105*, 477–490.

Zuberbühler, K. (2000a). Causal cognition in a non-human primate: field playback experiments with Diana monkeys. Cognition *76*, 195–207.

Zuberbühler, K. (2000b). Referential labelling in Diana monkeys. Anim. Behav. *59*, 917–927.

Zuberbühler, K. (2000c). Causal knowledge of predators' behaviour in wild Diana monkeys. Anim. Behav. *59*, 209–220.

Zuberbühler, K. (2001). Predator-specific alarm calls in Campbell's monkeys, *Cercopithecus campbelli*. Behav. Ecol. Sociobiol. *50*, 414–422.

Zuberbühler, K., and Wittig, R.M. (2011). Field experiments with nonhuman primates: a tutorial. In Field and Laboratory Methods in Primatology: A Practical Guide., (Setchell JM, Curtis DJ), pp. 207–224.

Zuberbühler, K. (2019a). Evolutionary roads to syntax. Anim. Behav. *151*, 259–265.

Zuberbühler, K. (2019b). Syntax and compositionality in animal communication. Phil Trans R Roc Lond B.